BIOMETRIKA

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY
W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY
EGON S. PEARSON

IN CONSULTATION WITH

HARALD CRAMER	J. B. S. HALDANE
R. C. GEARY	G. M. MORANT
MAJOR GREENWOOD	JOHN WISHART

VOLUME XXXIII

1943—1946

ISSUED BY THE BIOMETRIKA OFFICE
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

CONTENTS OF VOLUME XXXIII

Memoirs

PAGES

Miscellanea

First printed in Great Britain at the University Press, Cambridge
Reprinted by litho-offset
by the University Press, Oxford

Vol. XXXIII. Part I

April, 1943

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

ASSISTED BY

EGON S. PEARSON

Reprinted by offset-litho

ISSUED BY THE BIOMETRIC LABORATORY
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

VOLUME XXXIII, PART I APRIL 1943

MEDICAL STATISTICS FROM GRAUNT TO FARR
(Concluded*)

BY MAJOR GREENWOOD

VI. SOME ENGLISH MEDICAL STATISTICIANS IN THE EIGHTEENTH CENTURY

AFTER Petty more than fifty years passed before another Fellow of the College of Physicians took an interest in statistics, and he, if less eminent in political arithmetic, was much more eminent in the art of medicine; he was the elder Heberden. Heberden published no statistical work over his name, but there seems no reason to doubt the accuracy of his son's statement that the quarto volume containing a collection of the yearly Bills of Mortality in London from 1657 to 1758 and various essays was financed by Dr William Heberden and that he wrote the preface. This preface, an essay of 15 pages which ends rather abruptly, could hardly have been written by a layman. The following passage illustrates my remark:

> The deaths imputed to the measles are very remarkably different in different years; and yet it is possible that this disease is not in reality so very irregularly epidemical or fatal, as by the bills it appears to be. The scarlet fever and malignant sore throat often occasion such appearances upon the skin, as may easily be mistaken for the measles by better judges than the mothers and nurses, who thinking themselves able to distinguish this distemper, and equal to the management of it, often call in no other assistance. This mistake is well known to have been sometimes made within these few years, during which the scarlet fever and malignant sore throat have been so common. It may perhaps have happened in every year, in which an extraordinary number of deaths are charged to the measles: and consequently those two formidable distempers, (if they are two distinct distempers, and not one and the same) being disguised under the name of the measles, may have been older, and more general than is usually imagined.

The writer's observations upon the disappearance of plague have also something of a professional air—the fact that they are decidedly confused is no argument to the contrary. Sydenham taught (*Obs. Med.* **2**, 2) that plague depended upon (*a*) a special disposition of the atmosphere, (*b*) the transmission of an infecting matter, and held (*a*) to be primary, i.e. that without the atmospheric constitution there would be no epidemic. Heberden thinks the decline of the plague was due to the rebuilding of the city and—'probably the most effective'—the great quantity of water from the Thames and the New River, 'which, for the last century, has washed the houses so plentifully, and afterwards running down into the kennels and common sewers, constantly hinders, or weakens the tendency to putrefaction'. Heberden, unlike Sydenham, who

* The earlier sections were printed in *Biometrika*, **32**, 101–27 and 203–25.

believed the secret of the atmospheric constitution beyond the wit of man, seems to have attributed it to 'putrefaction', but, like Sydenham, attributed more importance to the atmospheric than the infective factor.

For the rest, Heberden continued Graunt's criticism of the material. In particular he gives good reason for thinking that beyond the omission of dissenters' christenings and burials, an important error arises from a balance of outward burials, that more coffins are taken out of London to be buried than are brought in. From an enumeration in Westminster he concludes that the deaths within the Bills are 20 % too few. He comments on an apparent increase in certain forms of death, such as apoplexies, lethargies and palsies. 'The practice of drinking spirituous liquors must, probably, answer for some part of this: and it might be of public use, if some attention were paid to the finding out of the other causes.' Rather optimistically, he thinks abundant amends might be made for these increases by the control of smallpox through inoculation. Upon this he makes a comment which has a very modern ring. 'For while inoculation prevails only among a part of any number of people, who all have an intercourse with one another, it may occasion as many deaths by spreading the distemper in, as it is called, the natural way, as it prevents among those, on whom it is practised.'

The volume contains a reprint of Graunt's work, of one of Petty's papers, and a new essay by Corbyn Morris. This essay, which shows signs of improving statistical technique, is not of medical interest except in its collection of deaths in age groups—an operation rendered possible by the introduction, in 1727, of an age classification of all deaths. Twelve age groups were giver. In Heberden's preface the importance of classifying by cause of death and age is emphasized.

From this material a life table was calculated by a fellow of the Royal Society named Postlethwaite (at the request of Heberden). This table, based upon deaths alone, is, for reasons already stated, of little value.

I doubt whether even the relative mortalities are correctly shown. The age distribution of the Bills (after 1726) for deaths was: under 2, 2–5, 5–10 and thence forward by decennia. Consequently one must distribute the deaths into single years of life upon some principle of interpolation. Neither J.P. (Postlethwaite) nor J.S. (Smart)—who made a life table for the first ten of the thirty years—states the principle on which he worked. But J.P. assigns 250 of the 363 deaths at ages under 2 to the first year of life and J.S. 290 of the 386 he had to manipulate, respectively 68·9 and 75·1 % of the total mortality under 2. Both ratios are much less than given by Halley's table, 83·3 %, which itself agrees admirably with the latest English population table (E.L. No. 10, Males) 83·5 %. If we applied the 83 % ratio it would raise the rate of infant mortality to 301 per 1000. Taking the figures simply as they stand, the survivors to age 6 years are fewer than Graunt estimated—54 % not 64 % survive.

Although, as has been said, the arithmetical values are very suspect, the

indication they give may be towards the truth. Creighton gave good reasons for concluding that London after the extinction of the Plague was less, not more, healthy. These reflexions are not without importance for they help to explain a certain fatalism, a scepticism as to the possibility of reducing the death-rate, which is noticeable in both statistical and medical literature for some time to come.

The next writer to be noticed is Thomas Short. Of this industrious investigator even Sir Norman Moore could obtain few personal particulars. He may have been born in 1690 and he died in 1772. He practised in Sheffield and was a Doctor of Medicine of a Scottish university but not a licentiate of the College. His principal works, *A General Chronological History of the Air, Weather, Seasons, Meteors, etc.*, published in 1749, and *New Observations, Natural, Moral, Civil, Political and Medical on City, Town and Country Bills of Mortality*, published in 1750, are differently assessed by the greatest historian of British medicine. Creighton pronounces the former to be rubbish but gives Short a not very hard pat on the back for the latter. 'That so much statistical or arithmetical zeal and exhaustiveness (in the work of 1750) should go with so total a deficiency of the critical and historical sense (in the work of 1749) is noteworthy, and perhaps not unparalleled in modern times' (Creighton, *Hist. of Epidem. in Britain*, 1, 405). Creighton's not very wide intellectual charity did not embrace statisticians.

It must be admitted that Short is decidedly *not* a writer to commend himself to an orderly minded, careful scholar from Aberdeen. Had he lived a century later we might have supposed that his literary model was Mrs Nickleby—he just runs on and on. A table (which must have been most troublesome to compile) of monthly christenings and weddings in various towns (in three extending over more than 150 years) leads him from arithmetical comparisons of the months most apt for procreation to a vigorous denunciation of luxury, of polygamy, of taxing common necessities, of the sale of army commissions, of novel reading, boarding schools and much else. But, although few if any would be able to read Short straight through without a rest, a good many less entertaining books might be included in a bed-side book case.

There is scarcely anything within the range of human interests upon which Dr Short has not something to say. On the whole he took a gloomy view of modern life in general and of his faculty in particular, and remarked that the 'improvements in surgery in general, have far out-stripped those in physick'. Surgeons he found to have generally less learning than physicians, but compensated for this by a closer application to the study of their own profession 'without jumbling the finite mind, and mixing studies of a different nature from their own, as of the dramatists, poets, classics, architecture, politics, history critics, logics, etc. They are also less liable to theories and false reasonings, have not that contempt of the ancients, nor of observations built on practice, improved

and directed by the understanding, and raised to the pitch of truth by a long enquiry into the effects of diseases and medicines.'

Short began his book with a clear plan—long before he had finished it the plan became an inextricable confusion. He argued that a statistical measure of health could not be obtained from the data of towns in general and the capital in particular, partly owing to the inaccuracy of the data, partly owing to the fact that towns attracted newcomers and were not maintained by the balance of births and deaths within the community. So he collected material from country parishes (he also obtained data from towns but the country parishes were his prime object of study). His first set of data was a collection of transcripts of the registers of eighty-three parishes. About 60 % of these parishes were from Yorkshire (mainly in the neighbourhood of Sheffield) and Derbyshire, but some from as far afield as Devonshire. He set them out in two periods, the first ending before the Restoration, the second coming down to the third decade of the eighteenth century. He had another set of eighty-three parishes for which the data covered only the second period. He classified his parishes in accordance with the nature of soil, altitude, exposure, whether wooded or bare, wet or dry. Sometimes his data covered more than a century, rarely less than 20 years. He gives the total number of baptisms, of burials (sexes distinguished) and marriages and works out the various ratios, the ratio of baptisms to burials being his chief tool.

From these data he draws a great many conclusions; for instance, respecting the salubrity or insalubrity of different soils and exposures. Most of these conclusions, it may be remarked, are now part of the common stock of lay and, perhaps, professional belief. But whether Short's data were adequate to sustain the conclusions is another question.

We may begin by taking purely arithmetical points into consideration, viz. whether, *assuming* that the parishes or groups of parishes were fairly comparable and *assuming* that the ratio of baptisms to burials is a fair measure of healthiness, Short had large enough figures for his purpose. For instance, two of his conclusions were that dry open sites of moderate elevation were healthier than a clay soil. I picked out of his list nine parishes of the former and five of the latter class. In his first (pre-Restoration) period, the parishes on dry, open sites had registered 4349 baptisms and 2644 burials, a ratio of 1·64. The five parishes on clay had 2875 baptisms and 1920 burials, a ratio of 1·497. So, as he said, the dry open sites give a higher ratio. But what is the order of magnitude of the error of sampling? We may safely hold that the standard error of the number of baptisms or burials is of the order of the square root of the observed number, or the ratio of standard error to number is of order $1/n^{\frac{1}{2}}$. From this we infer that the standard error of the ratio n_1/n_2 is given by n_1/n_2 times $(1/n_1+1/n_2-2r_{n_1n_2}.1/(n_1n_2)^{\frac{1}{2}})^{\frac{1}{2}}$. Clearly, the correlation between numerator and denominator must be large, so that the second factor lies between $(1/n_1+1/n_2)^{\frac{1}{2}}$

and $(1/n_1^{\frac{1}{2}} - 1/n_2^{\frac{1}{2}})$ and will be much nearer the second value. In a sample of 1000 Registration Districts I studied many years ago, the correlation of births and deaths was 0·73 %, in our particular case the standard error of the difference between the two ratios will be likely to be not much more than 20 % of its value.

From the purely arithmetical angle, I should conclude that Short was justified in holding that his ratios did really differ significantly, as we say, from site to site. But is it fair to assume that (1) the ratio of baptisms to burials is a good index of healthiness, (2) that the comparisons are in *pari materia*? These are much more difficult questions.

So far as concerns modern experience, it is clear that the ratio of births to deaths does not give a useful index of the rate of mortality. I made an experiment on Short's lines. I took out a sample of fifty Registration Districts for the decennium 1901–10, chosen in the following way. (1) No districts with more than 10,000 births or fewer than 1000 deaths were taken. (2) Those with many institutional deaths were excluded. For each the ratio of births to deaths was calculated and the following table formed:*

Ratio of births to deaths	No. of districts	Mean of standardized death-rates
2·0–	4	10·36
1·9–	6	10·88
1·8–	7	11·04
1·7–	6	11·60
1·6–	15	11·15
1·5–	7	10·72
1·4–	4	11·48
1·3–	1	13·01

It is true that the district with highest ratio has the lowest death-rate and the district with lowest ratio the highest death-rate, but in detail there is but little correspondence. Testing the same districts on the data of 30 years earlier, 1871–80, the same result appears. It would be very rash to conclude that because a district has a ratio of births to deaths above the average its standardized death-rate is below the average.

There are many reasons why a ratio of births to deaths may be a bad measure

* The districts selected were: Hambledon, Malling, Faversham, Romney Marsh, Uckfield, New Forest, Romsey, Hartley Wintry, Royston, Winslow, Witney, Oundle, St Ives, Caxton, Whittlesea, Lexden, Risbridge, Mildenhall, Bosmere, Plomesgate, Flegg, Cricklade, Melksham, Amesbury, Sturminster, Kingsbridge, Stratton, St Columb, Langport, Dursley, Ledbury, Wem, Mastley, Meriden, Shipston on Stour, Lutterworth, Spilsby, Hayfield, Garstang, Settle, Pateley Bridge, Gt Ouseburn, Saddleworth, Thorne, Pocklington, Skirlaugh, Easingwold, Bedale, Weardale, Brampton.

VII. SOME REPRESENTATIVE CONTINENTAL DEMOGRAPHERS OF THE EIGHTEENTH CENTURY

My object is to sketch the history of distinctively medical statistics in our own country; I have neither the knowledge nor, perhaps, the desire to cover a wider field; but it would be too insular entirely to neglect continental research contemporaneous with that described in the preceding section. I propose to discuss the work of some foreign writers which is relevant to that of the British authors mentioned in the preceding section. The most eminent contemporaries of Short were Deparcieux, Wargentin, Struyck, Kerseboom and Süssmilch, and of these Deparcieux, Struyck and Süssmilch are, I think, the most interesting, a Frenchman, a Hollander and a German. None of them was a physician. Deparcieux and Struyck were competent mathematicians. Struyck wrote on the general theory of probability, Süssmilch had no more mathematics than Graunt; but, of the three, Süssmilch is better known to posterity because he is frequently cited in books which circulate outside professional statistical circles. Deparcieux (1703–68) is the least voluminous and most attractive of the three. He published in 1746 a quarto of 132 pages (with tables) entitled *Essai sur les Probabilités de la Durée de la Vie humaine*, to which he added, 14 years later, a short appendix, and his book is a model of clear writing.

Deparcieux was fully alive to the dangers of basing a life table upon data of mortality alone, and was the first writer to construct what we should regard now (subject to a few reservations) as correct tables. Of course, like his contemporaries, he could not make bricks without straw, and no more than they could provide a general population life table. He had to use data which were not random samples of human experience and is careful to point this out. His new material was drawn from two sources, the data of tontines and the mortality experience of religious orders.

A tontine (the name is derived from that of the inventor Lorenzo Tonti, a Neapolitan banker) was a system of selling annuities on the following plan. The participants are formed into age classes, each entrant pays a capital sum and receives an annuity; as the annuitants die out the amount payable to the survivors is increased and the last survivor will enjoy an income equal to that distributed originally over all members of the age class. This was the general plan of a simple tontine (*The Wrong Box* will have made us familiar with a different application); there were various modifications, but in all an exact record of deaths at ages was essential.

Deparcieux used the data of tontines established in 1689 and 1696. He had to face many difficulties. In the first place, the tontines had a series of classes, one for those entrants under the age of 5, the next for lives from 5 to 10, and so on. What is the mean age of the members of each class? There would, as Deparcieux points out, be a bias in the first class (that of children under 5) in

favour of ages beyond the mean, because parents needed no statistics to convince them that the rate of mortality in the first and second years of life is higher than in the third and fourth or fifth. In the later classes, on the other hand, the bias would be in favour of entering at an age below the mean of the class limits. He makes a rather modest allowance for these factors by taking the age at entrance in the first class as 3 years, i.e. half a year more than the mean of the class limits, and in the next (and subsequent classes) as half a year less than the mean. The next difficulty is that his observations end in 1742, consequently rates of mortality at ages are derived from persons whose dates of birth are widely separated. Thus no members of the first class of the 1689 tontine can have been exposed to the risk of dying at ages beyond 57 (actually of 202 entrants, 105 were still living at the close of the observations). So a table obtained by welding these observations ignores any secular trend of mortality. It also ignores what, in modern assurance practice, is an important factor, viz. selection. A life aged n years is less likely to end within the year of entrance than a life of the same age entered 10 years earlier. In ordinary practice there are two reasons, self-selection and medical examination. In annuitant experience only the former is involved, but this is not the *less* important of the two.

In the discussion of this subject which will be found in Elderton and Oakley's *The Mortality of Annuitants* 1900–1920 (published on behalf of the Institute of Actuaries in 1924), the conclusion is reached that when *contemporaneous* lives are in question, this selection only operates seriously on the first year of annuitant life; for that year the rate of mortality is about 63 % of that suffered by annuitants of the same age who had purchased annuities 5 years earlier (what is called ultimate mortality). If then there were no secular improvement of mortality rates—as there has been over the last 60 years—and if there were no secular change in the social or economic class of annuitants, while we should expect a lighter mortality upon recent entrants, if, as in Deparcieux's data, we are only given survivors at quinquennial intervals, we should not expect large differences. Actually one can test a particular age group, viz. 45–50 on numerically extensive material. The 1689 tontine provides ten and the 1696 tontine nine groups of persons of this age the survivors of which 5 years later are recorded. It will be seen from Table 1 that the 634 'new' entrants in the 45–50 tontine class of 1689 suffered rather heavier mortality than the 118 survivors to that age from the youngest class. This, however, is merely picking out a single pair. The correct test is to treat the data together and inquire whether the hypothesis that the whole set of deaths and survivorships might have arisen by sampling a population for which the chance of living 5 years was simply the ratio of total survivors to total exposed, viz. $5009/5394 = 0{\cdot}9286$. Applying the appropriate test, viz. that known as the χ^2 test (with 18 degrees of freedom), one reaches $P = 0{\cdot}0346$. This is not a very improbable freak of chance. Compared with modern annuitants, these tontiniers of 200 years ago had a rate

of mortality some 40% greater than the annuitants of 1900–20 between the ages of 47 and 52.

Finally, one has the class of society from which annuitants are drawn. Deparcieux was of opinion that annuitants were mainly drawn (op. cit. p. 62) from the middle class of society 'ce sont les bons Bourgeois qui tiennent un honnête milieu entre toutes ces extremités, qui se font des Rentes viagères; et ce sont ceux-là qui deviennent ordinairement vieux'. Hence he judged that the rate of mortality suffered would be less than that of the general population.

Table 1. *Deparcieux's observations*

Tontine class	Exposed to risk at age 47	Survivors to age 52		Deaths	
		Observed	Expected	Observed	Expected
(1689 tontine)					
– 5	118	109	110	9	8
–10	181	173	168	8	13
–15	211	192	196	19	15
–20	216	196	201	20	15
–25	201	189	187	12	14
–30	263	249	244	14	19
–35	526	479	488	47	38
–40	472	440	438	32	34
–45	770	723	715	47	55
–50	634	575	589	59	45
(1696 tontine)					
5–10	134	130	124	4	10
–15	131	118	122	13	9
–20	108	103	100	5	8
–25	102	92	95	10	7
–30	147	135	137	12	10
–35	211	204	196	7	15
–40	220	200	204	20	16
–45	444	415	412	29	32
–50	305	287	283	18	22
	5394	5009		385	

The next part of his investigation related to the mortality experience of members of monastic orders. These he utilized with the same good sense and care.

In Table 2 are his l_x values, to which I have added those for English Life Table No. 9 Males (general mortality of 1930–2). The reason for putting l_{20} equal to 814 is simply that in his table for tontines where the starting point is age 3, his survivors to age 20 from an initial 1000 were 814.

The column headed Benedictines (*a*) is a methodologically correct table, viz. based on entrants followed until death, Benedictines (*b*) assumes a stationary

population and is not therefore so exact although it utilizes more data. Actually both tables give virtually the same results. It will be seen that to age 50 all the tables agree well; after age 50 the monks fare worse than the members of tontines and worse than the nuns. All have much worse mortality than the unselected general population of England and Wales 200 years later. Deparcieux attributes to selection the equality of tontine and monastic mortalities at younger ages and to the privations and austerities of the religious life a higher mortality at later ages.

In an investigation made by Dr S. Monckton Copeman and myself some years ago (*Report on Public Health and Medical Subjects*, no. 36, H.M.S.O. 1926) into the alleged low mortality from cancer of members of certain religious orders, we had occasion to study the general mortality experience. The result was that the mortality at ages over 25 of monks was rather more favourable than that of

Table 2. *Deparcieux's observations*

Age	Tontines	Survivors from age 20		Nuns	E.L. No. 9 Males
		Benedictines (*a*)	Benedictines (*b*)		
20	814	814	814	814	814
30	734	756	749	751	788
40	657	675	681	676	755
50	581	575	583	587	698
60	463	423	432	462	594
70	310	236	235	286	405
80	118	55	51	103	151

annuitants, that of nuns less favourable. The data were, however, scanty (monks 65 observed against 79·4 expected deaths; nuns 152 observed against 124·7 expected deaths).

Deparcieux has a few remarks on general medical-statistical questions (for instance, he urges strongly the importance of mothers nursing their infants), but nothing of much significance.

The statistical writings of Nicholas Struyck (1687–1769) are more voluminous than those of Deparcieux* and cover a wider field. Struyck was the son of an Amsterdam burgher and is said to have been in relatively easy circumstances. He enjoyed a considerable reputation as a writer on mathematical, statistical, geographical and astronomical subjects and was admitted a fellow of the Royal Society of London in 1749.

* They were collected and published in French translation at the instance of the Netherlands Assurance Society in 1912: *Les Œuvres de Nicolas Struyck, qui rapportent au calcul des chances*, etc., traduites du Hollandais par J. A. Vollgraff, Amsterdam, 1912, pp. 430.

Struyck was evidently a competent mathematician and also an industrious field worker who carried out or inspired in the Netherlands many town and village enumerations of population and vital statistical records. Like Deparcieux, he constructed life tables from annuitants' data and he certainly understood the correct arithmetical procedure. His data were, however, much fewer and he does not give sufficient details of his methods of interpolation and approximation to central ages for it to be possible to say precisely how he reached the life tables for males and females printed on p. 231 of his book. The original data were 794 males and 876 females (annuitants) observed for various periods and classified in quinquennial age groups. One has the impression that, although Struyck was a mathematician, he was not very sensitive to the dangers of basing conclusions upon small absolute numbers, and in his discussion of the vital statistics of London (op. cit. pp. 348–51) he has hardly given enough weight to the disturbing influence of migration and is perilously near the fallacy of a stationary population.

From the point of view of the medical statistician, Struyck is not a very suggestive writer. As demographer, we might rank him as technically superior to Short but medically less interesting. Like his contemporaries he can chase phantom hares in a thoroughly entertaining way. His finest example is in a section on multiple births. After a sober statistical inquiry he concludes that a case of quintuplets might reasonably be expected to occur sometimes in populations of the sizes of those of France and Germany—'it would be a very rare but not an incredible event'.

The case of the countess of Hennenberg, alleged to have brought to birth 364 or 365 infants simultaneously, does, however, strike him as 'absolutely fantastic and contrary to nature', and he carefully examines the legend. The statement was that the prolific mother produced as many children as the days of the year, and that the boys were named John and the girls Elizabeth. As Struyck justly observes it would be silly to have 182 Johns and 182 Elizabeths, and by careful research he arrives at a simple rational solution. The lady performed her feat on 26 March 1266; at that period the year began with the Feast of the Annunciation which was 25 March. So the birthday was the *second* day of the year and probably the mother had twins, one christened John the other Elizabeth. *Simplex munditiis!*

The name of Johann Peter Süssmilch (1707–67) is far better known than those of Deparcieux and Struyck although it is doubtful whether his *book* is often read. The perusal of 1201 pages of text and 207 of tables (the contents of the third edition of Süssmilch's book, published in 1765) requires a powerful appetite. If Süssmilch's literary style has less complexity than that of successors who wrote after German had become a 'literary' language, it has not much charm and few of us love propaganda. Süssmilch is a pure propagandist; the title of his book is: 'Die göttliche Ordnung in den Veränderungen des

menschlichen Geschlechts, aus der Geburt, dem Tode und der Fortpflanzung desselben *erwiesen*' (italics mine), von Johann Peter Süssmilch. He sets out to reveal the divine machinery for fulfilling the command: 'Be fruitful, and multiply, and replenish the earth, and subdue it.' The reason why his book has more interest for a statistician than, say, Warburton's *Divine Legation*, is that Süssmilch conceived the notion that vital statistics might be pressed into the service of orthodox Lutheran theology, and the diligence with which he pursued his arithmetical investigations gave his book importance. It was indeed the quarry from which Malthus obtained material when the interest aroused by the first edition of his famous *Essay* led him to expand what had been not much more than a Shavian paradox into a serious treatise.

As a demographer and statistician, Süssmilch was technically inferior to either Deparcieux or Struyck and, of course, far below Halley. He had none of Graunt's originality and made no methodological advance. But he was very industrious. He assembled not only a large collection of German data, similar to but wider than those of Short, but collected foreign material—including that of Graunt, King and Short—and his tables are of real value.

The general conclusions he reached—constancy of the sex ratio, greater mortality of towns, etc.—differ in no important respect from those of his predecessors or English contemporaries. His own life table (which gave an expectation of life at birth of 28·43 years) is constructed on the incorrect principles adopted by most of his contemporaries. He was, indeed, aware that to make a life table by summing the deaths at ages occurring in an increasing population was wrong and that the population he used was increasing, but he did not know how to do better—indeed, he had no material for doing better.

His contribution to purely medical statistics is small. He has a chapter on the statistics of causes of death and compares the distribution by causes in the London Bills 1728–57 with those for Berlin in the years 1745, 1750 and 1757. When allowance is made for differences of nomenclature and misprints, the proportional distributions by causes are not very different. He makes the sensible suggestion that if the Latin names of the diseases were given by the medical attendants in official returns international comparison would be facilitated. For the rest, his medical importance is slight. To criticize or make fun of his triumphant justification of the ways of God to man would be sorry trifling. Although a dull writer, he inspires a certain affection. He was a sincere, diligent man and in polemics more courteous than most. He may, perhaps, quite contrary to his intention, have had a rather depressing influence upon enthusiastic readers, in that he had no expectation of a great reduction of mortality rates and has often anticipated ideas which we usually attribute to Malthus. He perceived that at the current rate of growth the earth must eventually be overpopulated, but he argued that as the density of population increased the age of marriage would rise and consequently the fertility rate would decline. 'If,

however, fertility remained the same, it would only be necessary for the rate of mortality to increase a little, so that, as in large towns, one in 25 died' (op. cit. **1**, 267).

He devotes a whole chapter to what Malthus would call positive checks upon population and clearly does not expect these to be eliminated although, for the reason just quoted, he does not think plagues and wars essential conditions. One, perhaps only one, item of the vital statistical system gave the good man some qualms. His arithmetic leads him to conclude that in cities half those born are dead by the 20th year of life, and even in the virtuous country districts half are dead before the age of 25. 'What is the reason that God permits half to die before they can be of service to God and the world? All the labour and effort of birth and rearing seem to have been in vain' (op. cit. **2**, 312). In a worldly sense there is, he confesses, no explanation. One must think of earthly life as but a preparation for the hereafter.

The trend of this reasoning is not encouraging to the social or hygienic reformer. Perhaps Süssmilch did contribute a little to the view that not much could be done to reduce the general death-rate, that, at the best, town death-rates might be slightly improved. But I doubt whether he had much influence upon medical opinion in England. Statistics are not even now a favourite study of the medical profession; 200 years ago a voluminous German writer on vital statistics would have found very few readers in the College of Physicians.

VIII. METHODOLOGICAL ADVANCES

The writers who were the subject of the last sections all flourished in the first half of the eighteenth century and all have a claim to be reckoned as pioneers. Deparcieux and Struyck made definite contributions to the mathematical or arithmetical technique of life-table construction; Süssmilch and Short followed the path blazed by Graunt, but they explored a good deal of country, and Short, at least, had novel ideas as to the utilization of local records.

In the later years of the century various medical writers, for instance, Heysham, Haygarth and Percival, made effective use of local enumerations of population in their efforts to secure sanitary improvements. The public has no passion for statistics, still a death-rate *is* more telling than a mere enumeration of deaths. But none of these writers contributed anything new to statistical methodology, and simple arithmetic, not to speak of the labour of making unofficial counts of population, is not every man's hobby. Had the proposal for making an official census in the middle of the century been accepted, no doubt interest in political or medical arithmetic would have revived, but it did not pass the House of Lords. The only official data of large dimensions were still the London Bills. These were sometimes the subject of medical statistical comment. In 1800, the younger Heberden wrote a monograph the title of which suggested competition with Short or even Graunt. But it was not a successful venture and

is only remembered now (if at all) because of a statistical 'howler' which the iconoclastic Charles Creighton exposed with a satisfaction not melancholy.*

A typical example of the attitude of the better class of physicians towards statistics at the end of the eighteenth century will be found in *Observations Medical and Political on the Small-Pox...and on the Mortality of Mankind at every Age in City and Country...*, by W. Black, M.D., the second edition of which appeared in 1781. Dr Black, a medical graduate of Leyden and a licentiate of the College, who survived to 1829, reprinted the life tables of his predecessors. He was alive to the importance of the statistical method and its neglect ('In the course of many years' attendance upon medical lectures, in different universities, I never once heard the bills of mortality mentioned', op. cit. p. 119) and held that 'the detached observations of physicians or other literary individuals, confined perhaps to a small town or parish: a meagre detail of village remarks (*sic*), afford in many instances a foundation too slight to erect upon them any general or permanent conclusions' (op. cit. p. 119). He accordingly devoted most of his attention to the London Bills, which he subjected to a severe but cogent criticism, and set out in detail a sensible plan for the compilation of data in London by salaried officials with medical knowledge, which, had it been adopted, would have antedated the establishment of effective registration in London by more than 50 years.

One might explain the stagnation of medical statistical research by saying that there was not enough straw for ordinary brick makers to be employed, and no medical man of sufficient ingenuity (or temerity) to find a substitute for straw emerged. If that eminent fellow and, for a very short space of time, president of the College James Jurin had lived in the second instead of the first half of the eighteenth century, it is possible that the history of medical statistics would have been different, because, some years after his death, two famous mathematicians tackled a problem in which Jurin had taken keen interest and, as he himself was an accomplished mathematician, their method would have given him intellectual pleasure.

Jurin was an enthusiastic supporter of the practice of smallpox inoculation and wished to provide an adequate statistical proof of its value. Monk provides an eulogistic, Creighton a depreciatory account of what Jurin did. A fuller account is given by Miss Karn (M. N. Karn, *Ann. Eugen.* 4 (1931), 279 et seq.).

That Jurin proved the fatality of inoculated smallpox to be very much less than that of the natural smallpox, even Creighton admitted. But that he did much more can hardly be claimed. Jurin virtually assumed that inoculated smallpox did confer an immunity, on the basis of others' testimony and the famous experiment on six criminals, or rather on the one criminal who after

* Creighton, *History of Epidemics in Britain*, 2, 747. Heberden made two mistakes: (1) He did not recognize that 'Griping in the Guts' of the Bills of Mortality was mainly the Diarrhœa of young children. (2) That a gradual transfer from this heading to that of 'Convulsions' had been going on.

inoculation was deliberately exposed to natural infection (see Creighton, op. cit. p. 480). Whether Jurin deserves to be sneered at because he did not do what was impossible, or whether the assumptions he made were unreasonable, are questions I shall not discuss. The mathematicians added nothing to the biological discussion, the interest of their work is purely intellectual, viz. by showing how to make the most of imperfect material. The problem proposed by Daniel Bernoulli was this.

Let us assume that inoculation completely protects against dying from smallpox and that those who are thus saved from the smallpox are neither more nor less likely to die of other causes than persons who never take smallpox, then what would be the effect on general mortality of the total eradication of smallpox? Put more picturesquely, how many years would be added to the average span of human life if smallpox were extinct?

In modern times, questions like this have often been put and answered, because we know with fair accuracy the numbers living by sex and age and the numbers dying from different causes also by sex and age. In the famous Supplement to the *35th Annual Report of the Registrar-General*, Farr dealt with several causes. His method was simple. He subtracted from the central death-rate at any age due to all causes of death the central death-rate due to the special cause, and deduced from the resultant series of modified death-rates the appropriate life table constants. These he compared with those of the general life table. He found in this way that if phthisis were eliminated the expectation of life at birth (males) would be increased from 39·7 to 43·96 years. The elimination of the zymotic diseases would increase the mean lifetime to 46·77 years.

Farr was, of course, aware that the assumption, viz. if a particular cause of mortality was eliminated the death-rates from other diseases would not be affected, might not be justified—indeed, he had written with respect of Watt's lugubrious substitution theory, in accordance with which we gain little by eliminating one disease, its killing power will be taken by another. Farr's method is quite satisfactory as an arithmetical method but requires data not available in the eighteenth century. Bernoulli made two assumptions. The first that mortality rates from all causes were known (for his arithmetical calculations he used Halley's table although he did not quite correctly appreciate the meaning of Halley's phrase 'age current'), the second that the attack and fatality rates of smallpox were independent of age. He then reasoned thus:

Suppose there survive to age x by the life table P_x persons. Of these s, say, have not had smallpox; if $1/n$th of those who have not had smallpox were attacked within a year and $1/m$th of these die of smallpox, what is the value of s in terms of P_x, m and n? If dx is an element of time, $s\,dx/n$ will be attacked and $s\,dx/(mn)$ will die of smallpox within the element of time dx, and so there die

from other diseases $-dP_x - s\,dx/(mn)$ because $-dP_x$ is the total mortality. But we are only interested in s, so the decrement through mortality $-dP_x - s\,dx/(mn)$ must be multiplied by s/P_x, and we reach the equation

$$-ds = \frac{s\,dx}{n} - \frac{s}{P_x}\left(dP_x + \frac{s\,dx}{mn}\right),$$

the solution of which is
$$s = \frac{mP_x}{(m-1)\,e^{x/n} + 1}.$$

So s is known. Now let z be the number who would have survived to age x had there been no smallpox. Reasoning as before

$$-dz = -\frac{z}{P_x}\left(dP_x + \frac{s\,dx}{mn}\right).$$

The integral of which is
$$z = \frac{P_x m\, e^{x/n}}{(m-1)\,e^{x/n} + 1}.$$

This is the solution. Bernoulli put $n = m = 8$ and concluded that the elimination of smallpox would, on these assumptions, add about 3 years to the mean life-time.

D'Alembert criticized Bernoulli's assumption that m and n were constant and replaced his equation by the formally simpler equation

$$dz = \frac{z}{P_x}\,dP_x + \frac{z}{P_x}\,du,$$

where du is the increment of mortality in time dx due to smallpox. The formal solution is
$$z = P_x \exp\left[\int_0^x \frac{du}{P_x}\right]$$

Isaac Todhunter commented sub-acidly on this: 'The result is not of practical use because the value of the integral is not known. D'Alembert gives several formulae which involve this or similar unfinished integrations' (*History of the Theory of Probability*, p. 268). Todhunter's comment is just so far as concerns the situation when Bernoulli and D'Alembert wrote. If, in addition to a table of general mortality, one has knowledge of the deaths at ages due to smallpox, then by means of the theorem known as the Euler-Maclaurin expansion, it *is* possible to evaluate the integral and reach a solution on D'Alembert's lines as Miss Karn (op. cit. pp. 303 et seq.) has shown. But if we do have this information, the much less laborious method of Farr is adequate.

But that does not mean that the attempt of Bernoulli and D'Alembert was futile, a mere display of mathematical fireworks. The situation in which they found themselves recurs time and again in the history of statistics, indeed of all branches of science. Often a practical man objects that a mathematician will write down equations in general terms which cannot be solved and are therefore, as the practical man urges, of no use to him. Sometimes the practical man is

right, but not always; not even usually. Even when the equations cannot be solved, in the sense that certain 'constants' cannot be determined or certain integrals evaluated, methods of approximation, even inspired guesses, may lead to truth. Fifty years after Bernoulli and D'Alembert, E. E. Duvillard* published a monograph which, although seldom read, for it is scarce and 'practically' obsolete, has been rightly described by Farr as a classic of vital statistics. Duvillard set himself the same problem with the difference that vaccination was substituted for inoculation as the prophylactic, and this book, of nearly 200 quarto pages, may still be read with profit.

Duvillard lived before the days of Cauchy and mathematical rigour; no doubt much of his work would hardly satisfy the standard of a modern pure mathematician. Perhaps on that account it can be read by the amateur with comparative ease, and one may take hints of how to tackle problems for the solution of which complete statistical data are still to seek. There is no proverb the vital or medical statistician should more often repeat than the saying that the best is often the enemy of the good. It is no doubt foolish to suppose, as, according to Isaac Todhunter, Condorcet did suppose, that truth could be extracted from any data, however imperfect, provided one used formulae garnished with a sufficient number of signs of integration. It is more foolish to neglect even rough approximations to unattainable solutions. But, so far as concerns our predecessors in the College, indeed in the medical profession as a whole, the seed scattered by the foreign mathematicians fell upon stony ground. Between Short and Farr, no British physician made a contribution to statistical knowledge of much importance. I have spoken of the younger Heberden's brochure. William Woolcombe of Plymouth, in a tract on the alleged increase of tuberculosis, published in 1808, showed a better grasp of statistical method than the more famous physician.

The question Woolcombe examined was whether mortality from tuberculosis of the lungs were increasing. The statistical fact was that in well-kept registers he had examined the proportion of deaths assigned to consumption had certainly increased towards the end of the eighteenth century. Woolcombe was alive to the fact (often ignored by medical writers after his time) that the proportional mortality of a disease might increase although its absolute rate of mortality was stationary or even diminishing, and he tested his conclusions by a quite logical *ex absurdo* argument. Taking the assumption that at the beginning of the eighteenth century mortality was 1 in 36 and that the proportional mortality from phthisis was a third less than in 1801, he concluded that the general rate of mortality at the beginning of the nineteenth century must be as low as 1 in 54, unless the rate of mortality from phthisis had increased. But it was certain that in 1800 the general rate of mortality was higher than 1 in 54, at least 1 in 47. Reversing the process, viz. assuming the rate in 1801 to be known, the con-

* *Analyse et tableaux de l'influence de la petite vérole sur la mortalité à chaque âge*, Paris, 1806.

clusion was reached that the rate of mortality at the beginning of the eighteenth century must have been 1 in 27 unless the rate of mortality from phthisis had increased. This Woolcombe thought improbably high. He may have been wrong, but his method was rational. That was the best piece of medical statistical reasoning I have found in English medical literature between Short and Farr.

In 1800 the taking of a census was authorized by the legislature and not a government department, but the Speaker of the House of Commons was charged with the responsibility. Naturally, Mr Speaker passed over the actual work to one of his subordinates, and fortunately that subordinate, John Rickman, whose name is immortalized by the fact that he was a friend and correspondent of Charles Lamb, was really interested in statistics. In the report on the enumeration of 1801 comments are scanty, but they increased in subsequent volumes. Rickman was wholly responsible for the work down to the report of 1831 and, although he made no advance in statistical method, he did valuable work, particularly in calling attention to the high rate of mortality in the industrial north-west and in estimating past populations of the country. But Rickman was not professionally interested in medical questions, and before Farr no medical man utilized the new material effectively. As will be seen in the next section, the first English writer to publish a work under the title Medical Statistics was rather old fashioned in his treatment of the subject.

IX. THE END OF AN EPOCH

Almost at the end of the period I have chosen was published the first English book specifically devoted to Medical Statistics, *Elements of Medical Statistics*, by F. Bisset Hawkins, printed in 1829. It is a slender volume of 233 pages similar in format and size to the *Principles of Medical Statistics* published a little more than a century later, in 1937, by my friend and colleague Dr A. Bradford Hill.

Hawkins's book was an expansion of the Gulstonian Lectures of 1828; its author's long and useful life connects men still living with what seems a remote past. He was born in 1796, and there are still more than a dozen fellows of the College who may have sat in Comitia with him. He was admitted a fellow on 22 December 1826 and died in 1894. The copy of his book which I have read was presented by him to the Statistical (now Royal Statistical) Society in 1834 and contains corrections in his hand. Hawkins defines the province of Medical Statistics to be 'the application of numbers to illustrate the natural history of man in health and disease'. In his numerical statements he uses three indices; the ordinary crude death-rate—always expressed as one death in such or such a number—the 'probable life', i.e. the age to which half these born attain; the 'mean life', i.e. the average age at death. He was certainly aware that the age and sex constitution of a group affects the death-rate. Thus (op. cit. p. 20) he writes: 'In discussing the mortality of manufacturing towns or districts, it is just to remark that the small proportion is not always *real*; because a constant

influx of *adults* is likely to render the number of deaths less considerable than that which could occur in a stationary population composed of all ages.' From the use of the term *stationary population* in this passage we may also, perhaps, infer that Hawkins knew the limitations of utility of such indices as mean age at death or *vie probable*, but I cannot fairly say that in making comparisons he calls attention to the dangers.

A modern treatise, such as that of Dr Hill, devotes a large space to methods of evaluating errors of sampling or, to speak loosely, the precautions to be taken when the observations are few in number and may not have been taken without bias. Some of the methods still employed had been invented by mathematicians before Hawkins's day, but he did not use them. On p. 32 we read: 'The annual mortality of Nice, though a small town, and enjoying a factitious reputation of salubrity, is 1 in 31; of Naples, is 1 in 28. Leghorn is more fortunate, and sinks to 1 in 35. We instance those places as being the frequent resort of invalids; but how astonishing is the superiority of England, when we compare with these even our great manufacturing towns, such as Manchester, 1 in 74; such as even Birmingham, 1 in 43; or even this overgrown metropolis, where the deaths are only 1 in 40.'

In the copy I have read, the sentence 'such as Manchester, 1 in 74' has been struck through, apparently by the author. But, even with this emendation, the comparison, to the glory of our country, is, well, tendentious.

Indeed, one must admit, however regretfully, that Hawkins's book is uncritical. He has been diligent and brought together numerical data from all parts of the world and was certainly one of the first physicians to advocate a serious study of hospital records, but one can hardly say that, as a statistician, he was better equipped or more efficient than Dr Short in 1750. But his modesty is disarming: 'I should be amply rewarded if the present humble essay should form a temporary repository of the most important of their labours; if it should become one of the early milestones on a road which is comparatively new, rugged as yet and uninviting to the distant traveller, but which gradually discloses the most interesting prospects, and will at length, if I do not deceive myself by premature anticipation, largely recompense the patient adventurer' (op. cit. p. vii).

According to Munk (*Roll*, **3**, 304) Hawkins was instrumental in obtaining the insertion in the first Registration Act of a column containing the names of the diseases or causes by which death was occasioned. 'At first the insertion was voluntary; it has since been made compulsory; and has produced important additions to medical and statistical science through the indefatigable labours of Dr W. Farr.'

So the name of Francis Bisset Hawkins deserves a place in the roll of benefactors to medical statistical science.

Eight years after the publication of Hawkins's Gulstonians there appeared,

as Chapter IV of the fifth part of McCulloch's *Statistical Account of the British Empire* (**2**, 567–601, London, 1837), an article on 'Vital statistics; or the statistics of health, sickness, disease and death', the work of William Farr, then in his 30th year and still a general practitioner and free-lance medical journalist. It contains perhaps a quarter of the number of words in Hawkins's book and is not free from the quaint moralizing not always wholly relevant to the statistical theme which was characteristic of Farr, but it ranks not much below Graunt's 'Observations' as an original contribution to medical-statistical science.

Farr proposed to examine 'the mortality, the sickness, the endemics, the prevailing forms of disease, and the various ways in which, at all ages, its [The British Population's] successive generations perish'.

Slow as had been the progress of official statistics between 1662 and 1837, there had been progress. The four censuses of 1801–31 provided reasonably complete accounts of total populations. In 1821, information as to age was invited and eight-ninths of the population accepted the invitation. In 1831 the clergy were asked to return not merely totals of burials but burials classified by ages for the 18 years ending in 1830. These latter returns were incomplete, but it was possible for a lesser man than Farr to approximate to a statement of rates of mortality at ages at least for the period centring on 1821. To Farr's annoyance, the census takers of 1831 did not ask for the ages of the enumerated, contenting themselves with an enumeration of males under and over 20 years of age. The data for computing mortality rates were particularly defective for towns, but a few instances of quite good voluntary enumerations, e.g. for Carlisle and Glasgow, were available.

In handling national rates of mortality at ages, Farr's article does not display any conspicuous originality; he, quite properly, used the work of predecessors and he does not comment on the defects of the data. He does, however, call attention to particular rates of mortality, for instance, those of the troops, in an emphatic way. 'By the subjoined table of the mortality of the British army it will be seen that the soldier, in the prime of his physical powers, is rendered more liable to death every step he takes from his native climate, till at last the man of 28 years is subject, in the West Indies, to the same mortality as the man of 80 remaining in Britain.' According to his table, the average strength of British troops in Jamaica and Honduras between 1810 and 1828 was 2528; in the year of least mortality the rate 47 per 1000, the average 113 and the maximum 472! In the United Kingdom the average rate was 15 per 1000.

The most original part of Farr's essay is his treatment of sickness. Here national statistics were not available; more than 70 years were to pass before any nation-wide data were collected, and the statistics of morbidity still lag behind those of mortality. All Farr had were some data of benefit societies and returns relating to workers in the Royal Dockyards and employees of the East India Company. He begins by stating that in manhood for every death

we may reckon two persons constantly sick. It is not quite clear how he reached this ratio, but probably from a comparison of the mortality rates for 1815–30 shown in a table on p. 568 of his article with some theoretical rates deduced by Edmonds for Friendly Societies (op. cit. p. 574). One has:

Age	Sickness rate per 1000	Mortality rate per 1000
20–30	17·2	10·1
30–40	23·0	11·4
40–50	31·0	14·9
50–60	45·1	23·4
60–70	93·6	45·3

Taking the general rate of mortality to be 21·3 per 1000 and the population of England and Wales to be 14,000,000, he concludes that 600,000 persons are constantly sick and that the productive power of the community is reduced by one-seventeenth part (he has made allowance for attendance on the sick). He works out from the limited data available the relation between sick-time and age and concludes that it increases in geometrical progression up to the age of 50. He asks how much sickness exists among the labourers of the country independently of those definitely incapacitated by disease. Data for the Royal Dockyards lead him to conclude that 2% are constantly kept at home by illness.

In the last section of his article, Farr considers particular diseases. An instance of his acumen is to be seen in his criticism of the view (held in 1837 as in 1937) that insanity was on the increase. He pertinently remarks that if the less barbaric treatment of lunatics diminished the mortality rate a higher proportion of enumerated lunatics would be perfectly consistent with a steady rate of morbidity.

His data for rates of mortality by causes were scanty. For London over a long period he had causes of death in age groups and, from an estimate of total mortality in age groups, could pass back to rates at ages by causes. Heysham's Carlisle data were medically and statistically more precise but limited to one not large town. The data of the Equitable Assurance Society were numerous but, as, of course, Farr knew and emphasized, related to a select class of the population.

Some of his general conclusions were as follows:

It has been shown that external agents have as great an influence on the frequency of sickness as on its fatality; the obvious corollary is, that man has as much power to prevent as to cure disease. That prevention is better than cure, is a proverb; that it is as easy, the facts we had advanced establish. Yet medical men, the guardians of public health, never have their attention called to the prevention of sickness; it forms no part of their education. To promote health is apparently contrary to their interests: the public do not seek the

shield of medical art against disease, nor call the surgeon, till the arrows of death already rankle in the veins. This may be corrected by modifying the present system of medical education, and the manner of remunerating medical men.

Public health may be promoted by placing the medical institutions of the country on a liberal scientific basis; by medical societies co-operating to collect statistical observations; and by medical writers renouncing the notion that a science can be founded upon the limited experience of an individual. Practical medicine cannot be taught in books; the science of medicine cannot be acquired in the sick room. The healing art may likewise be promoted by encouraging post-mortem examinations of diseased parts; without which it is impossible to keep up in the body of the medical profession a clear knowledge of the internal change indicated by symptoms during life. The practitioner who never opens a dead body must commit innumerable, and sometimes fatal, errors (op. cit. p. 601).

Farr's article closes the epoch Graunt's book opened. The seventeenth-century pioneer did not live to see the ground he broke bear a crop. The high gods used Farr better; he lived to create the best official vital-statistics of the world. It is true that the lessons he taught were learned but slowly, either by the public or the profession. The *Annual Reports* of the Registrar-General will not be found among the frequently consulted volumes on the shelves of fellows of the College of Physicians. But something has been learned. The moral truism that human vanity is a deadly sin, now exemplified on a world-wide scale, is illustrated on the humbler scale of those topics which have been my life's work and the subject of these lectures. The distrust of 'mathematical' methods which is still general in our profession is not primarily due to the mere intellectual difficulty of learning 'mathematical' methods; much that all medical students must learn is at least as difficult.

The roots are deeper. They begin with the exaggerated claims of the iatro-mathematicians of the late seventeenth and early eighteenth centuries. The personal popularity of such men as Freind and Jurin did not conceal the fact that pathology and clinical medicine reduced to mechanical and quantitative theorems, and 'proofs' were of not much greater value in the treatment of sick men than skill in playing chess to the commander of an army. It is arguable that a talent for playing chess might, other things equal, be of advantage to a military strategist (Napoleon Bonaparte was very fond of chess and played so badly that it was difficult for his staff to avoid winning), but other things are not equal. In later times, when the intellectual prestige of mathematical science had grown enormously, it was observed that such an Admirable Crichton as our Thomas Young was inferior as a practical physician to many fellows of lesser fame. In our generation when the professional mathematicians who, 50 years ago, rather despised mere statistics, have increasingly devoted themselves to the improvement of the general theory, the complexity of statistical investigations has done little to attract the amateur, and intellectual modesty has not been the most conspicuous virtue of statistical authors. Perhaps, too, it is not easy for an experienced physician 'to renounce the notion that a science can be founded upon the limited experience of an individual'.

The moral I should draw from the history of medical statistics is that the intellectual courage of an amateur often succeeds where erudition fails. While even the purest of mathematicians would not claim that statistics is only a branch of mathematics, the hardiest contemner of algebra would admit that a training in mathematical method is an advantage to the practical statistician. The mathematician would surely agree that a knowledge of the material subjected to analysis was valuable, even if not so essential as a 'practical' man would claim.

Judged by contemporary intellectual standards, neither Graunt nor Farr was a mathematician; Graunt had no medical training, Farr's clinical experience was meagre. In respect neither of method nor subject-matter was either man an expert. But they both had intellectual curiosity and courage: one may say, if one pleases, the spurious courage of the man who is brave because he does not know what the dangers are. But, as Gilbert Chesterton once said, 'There is no real hope that has not once been a forlorn hope.' In graver matters than medical statistics and more than once in our national history salvation has been wrought by courageous amateurs who acted while professionals doubted.

Those who cannot disclaim a professional status in statistics, whether officials or professors, may learn a lesson from history. It is conveyed in the four words: *maxima debetur puero reverentia*, construing *puer* by amateur or beginner or enthusiast. It is weary work to read statistical 'proofs' of this or that aetiological theory of cancer, or proposals for this or that impossible statistical investigation. But it is treachery to science to rebuff any genuinely inquisitive person; the discovery of another Graunt in a shop or another Farr in the surgery of a general practitioner would repay the life-long boredom of all extant civil servants and professors of statistics.

A STUDY OF A SERIES OF HUMAN SKULLS FROM CASTLE HILL, SCARBOROUGH

By K. L. LITTLE, B.A.,* *The Duckworth Laboratory, University Museum of Archaeology and Ethnology, Cambridge*

1. Introduction

Much of our knowledge of the early settlers of north-east Yorkshire, and more particularly of its medieval inhabitants, is derived from the archaeological material provided by Scarborough Castle Hill. Here, during the years 1921–5, a notable series of excavations, carried out under the supervision of Mr F. G. Simpson, M.A., Hon. F.S.A. (Scot.), revealed at three quite distinct levels an early Iron Age village, a Roman signal station, and three chapels, one earlier and two later than the Norman conquest.

The Iron Age remains, derived from a prehistoric occupation-layer associated with a series of forty-two or more rubbish pits and archaeological material, are doubtless a relic of those immigrants who arrived about the seventh century B.C. and sought a temporary dwelling-place on the headland between the landing-places of Scarborough, the present North and South Bays. Whether these settlers moved northwards along the coast towards the jet-producing region, or whether they were few in number and shortly vanished from the countryside can only be surmised. The fact that Castle Hill is almost the most northerly site at which evidences have been observed of active immigration from the Continent at this period suggests that their number could not have been large, although it is possible that a more numerous Iron Age population lived in the neighbouring Wolds.

The subsequent Roman occupation of Britain lasted until the early part of the fifth century A.D. when Roman troops were finally withdrawn from the country. The signal station which the excavations on Castle Hill revealed appears to have been built, like a number of similar coastal forts, at the instigation of the Roman general Theodosius, as a small garrisoned outpost against the Saxon raiders, some half-century before the final evacuation. Racially, therefore, the presence of this small body of men could have had little effect on the indigenous inhabitants of the district.

Thereafter, in the period preceding the arrival of the Norman conquerors, the surrounding district appears to have provided a battle-ground for Danes and foraying Norwegians. In 1066, what was probably still a small settlement at the site of present-day Scarborough was pillaged and its chapel and wooden houses were fired by the men of Tostig and Harald Hardraada. It is probable that this attack, and subsequent wasting of the north by William, practically extinguished the early township, and that it did not recover until the second quarter of the twelfth century, when William le Gros of Albemarle chose the site on the promontory to lay the foundations of his castle.

The rebuilding of the chapel took place after the foundation of the castle. It seems that its history after that period is shared largely with that of the castle, and we may assume

* The writer wishes to acknowledge the very substantial assistance given by Dr Morant in the composition of this paper.

absorption of the alveolar ridge, indicative, presumably, of pyorrhoea. This feature did not appear to have affected the fixture of the relevant teeth to any particular degree. Protruding teeth were few, but in a small number of cases the incisors were prominently displayed owing to a slight degree of alveolar prognathism.

3. The variability of the Scarborough population

Standard deviations and coefficients of variation are given in Table 2 for the male series in the case of all characters for which thirty or more measurements are available. The numbers on which these constants are based (given in Table 3) range from thirty to forty-three. The female series is too short to give any estimate of variation worth considering. The male series may be supposed long enough to indicate any outstanding peculiarity of the population in the respect considered. It was only compared in detail with the Farringdon Street series of seventeenth-century London skulls (Hooke, 1926). This has often been used in estimating the relative variation of populations represented by cranial series, though it is rather less homogeneous than most of the others from British sites.

Table 2. *Standard deviations and coefficients of variation of the Scarborough male series (with probable errors)*

	S.D.	C. of V.
C*	94·3 ±7·2	6·23±0·48
L	8·05±0·59	4·34±0·32
B	5·50±0·41	3·77±0·28
B'	4·99±0·36	5·00±0·36
H'	5·22±0·38	3·92±0·29
S'_1	3·96±0·33	3·46±0·29
S'_2	6·87±0·57	6·03±0·50
S'_3	4·03±0·32	4·12±0·33
S_1	6·26±0·52	4·85±0·40
S_2	7·76±0·64	6·18±0·51
S_3	7·08±0·57	5·95±0·48
S	16·2 ±1·3	4·35±0·36

	S.D.	C. of V.
$\beta Q'$	11·8 ±0·87	3·73±0·27
U	16·7 ±1·2	3·17±0·24
fml	3·19±0·23	8·69±0·64
LB	4·21±0·33	4·18±0·33
GL	5·62±0·44	6·05±0·48
$G'H$	6·49±0·57	9·26±0·81
GB	4·71±0·40	5·03±0·42
NH, L	3·04±0·25	5·98±0·48
NB	1·66±0·13	6·67±0·51
O_1L	1·66±0·13	4·01±0·32
O_2L	2·92±0·24	8·46±0·69
G'_1	3·71±0·32	8·15±0·71

	S.D.
100 B/L	4·41±0·33
100 H'/L	4·13±0·30
100 B/H'	4·65±0·36
$Oc.I.$	4·06±0·33
100 $NB/NH, L$	3·26±0·27
100 $O_2/O_1, L$	7·10±0·57
$N\angle$	3°·56±0·31
$A\angle$	4°·70±0·41
$B\angle$	3°·67±0·32

* The constants are for reconstructed capacities.

Using coefficients of variation for absolute measurements and standard deviations for indices and angles, it is found that the Scarborough constant is the greater in the case of eighteen characters and the Farringdon Street is the greater in the case of the remaining fifteen. There are no markedly significant differences. The most significant are for C (Δ/P.E. $\Delta = 3{\cdot}9$), 100 $NB/NH, L$ (3·2) and GB (3·0), for which the Farringdon Street constant is the greater, and $Oc.I.$ (3·9), $G'H$ (3·4), G'_1 (3·1) and 100 $O_2/O_1, L$ (3·1), for which the Scarborough constant is the greater. Judging from all the characters, there appears to be no appreciable difference between the variabilities of the two populations. It may be noted that the standard deviation of the cephalic index for the Yorkshire series is unusually high, and it would have to be taken to indicate clear racial heterogeneity if found for a larger sample.

4. Comparisons between the Scarborough and other British series

Mean measurements of the Scarborough series are given in Table 3. The sex ratios of the absolute measurements are unexceptional, and in view of the small numbers of specimens the agreement between the corresponding male and female indices and angles is as close as

that expected for samples representing the same population. In such a case the mean female cephalic index is expected to be about one unit greater than the male index when the numbers are adequate, but no significance can be attached to the fact that a difference

Table 3. *Mean measurements of the Scarborough and two continental series of skulls**

	Scarborough		S.W. Norwegian‡	Belgian§
	Male	Female	Male	Male
C	1513·9 ± 10·2 (39)†	1365·0 (15)†	1467·5 (127)	—
L	185·5 ± 0·83 (43)	179·9 (16)	185·4 (144)	183·9 (133)
B	146·0 ± 0·59 (40)	139·9 (16)	143·6 (143)	145·3 (133)
B'	99·5 ± 0·51 (43)	97·4 (16)	97·9 (145)	98·2 (52)
H'	133·1 ± 0·54 (42)	126·7 (15)	129·3 (144)	131·1 (52)
S'_1	114·4 ± 0·46 (33)	113·1 (11)	112·5 (146)	—
S'_2	113·9 ± 0·81 (33)	112·4 (12)	112·0 (143)	—
S'_3	95·5 ± 0·46 (35)	96·9 (12)	96·6 (143)	—
S_1	129·0 ± 0·73 (33)	125·4 (11)	127·7 (146)	125·5 (52)
S_2	125·5 ± 0·91 (33)	122·7 (11)	124·8 (144)	127·3 (52)
S_3	118·9 ± 0·81 (35)	118·9 (12)	118·2 (143)	119·2 (52)
S	372·9 ± 1·9 (33)	366·2 (13)	370·2 (144)	372·1 (52)
$\beta Q'$	315·7 ± 1·2 (42)	302·5 (16)	308·6 (145)	—
U	525·9 ± 1·8 (41)	512·5 (16)	—	528·9 (52)
fml	36·7 ± 0·33 (43)	35·0 (16)	36·5 (139)	34·5 (33)
LB	100·6 ± 0·47 (36)	95·6 (14)	99·9 (142)	101·2 (33)
GL	92·8 ± 0·62 (37)	88·7 (14)	—	96·5 (33)
$G'H$	70·0 ± 0·80 (30)	63·4 (13)	72·2 (112)	—
GB	93·7 ± 0·56 (32)	87·0 (14)	94·7 (114)	—
J	133·2 (25)	122·4 (10)	134·0 (114)	132·0 (52)
NH, L	50·8 ± 0·35 (35)	47·8 (13)	52·5 (123)	—
NB	24·9 ± 0·18 (39)	23·4 (14)	24·6 (123)	23·5 (52)
O_1L	41·4 ± 0·19 (36)	41·0 (16)	41·3 (127)	—
O_2L	34·5 ± 0·33 (35)	35·7 (15)	34·5 (127)	33·6 (52)
G'_1	45·5 ± 0·46 (30)	43·7 (12)	46·0 (85)	—
G_2	40·6 (20)	35·6 (8)	41·3 (68)	—
$100\,B/L$	79·0 ± 0·47 (40)	77·6 (16)	77·6 (142)	79·0 (133)
$100\,H'/L$	71·8 ± 0·43 (42)	70·7 (15)	69·8 (141)	71·0 (52)
$100\,B/H'$	109·9 ± 0·50 (39)	110·3 (15)	{111·1 (143)}	{110·9 (52)}
$Oc.I.$	59·7 ± 0·46 (35)	58·5 (12)	{58·4 (143)}	—
$100\,G'H/GB$	74·7 (23)	74·8 (13)	{76·2 (112)}	—
$100\,NB/NH, L$	49·3 ± 0·39 (32)	49·4 (13)	47·0 (120)	—
$100\,O_2/O_1, L$	81·4 ± 0·81 (35)	87·8 (15)	82·2 (127)	—
$100\,G_2/G'_1$	91·1 (18)	81·6 (8)	90·6 (61)	—
$N\angle$	63°·5 ± 0·44 (30)	64°·4 (13)	—	—
$A\angle$	74°·7 ± 0·58 (30)	74°·1 (13)	—	—
$B\angle$	41°·8 ± 0·45 (30)	41°·4 (13)	—	—

* The mean indices in curled brackets were found from the means of the component lengths instead of from values for individual skulls.

† The capacities of the Scarborough skulls are estimates obtained by applying reconstruction formulae: see Appendix.

‡ Pooled means for the Bergen, Jaeren and Sogn series described by Schreiner (1939). Means for several other characters are provided in his monograph, and of these $fmb = 31{\cdot}2$ (142), and $100\,fmb/fml = 85{\cdot}8$ (139), are used in biometric practice in calculating coefficients of racial likeness.

§ Pooled means for series described by Heger & Dallemagne (1881). Other means are: Broca's $Q' = 311{\cdot}2$ (52), $NH' = 50{\cdot}1$ (52), $100\,NB/NH' = 47{\cdot}1$ (52), $O'_1 = 39{\cdot}0$ (52), $100\,O_2/O'_1 = 86{\cdot}1$ (52), $fmb = 29{\cdot}1$ (33), $100\,fmb/fml = \{84{\cdot}3\ (33)\}$.

of the opposite sign is noted here. The mean female orbital index is normally appreciably greater than the male.

The female series must be considered too short to use for comparative purposes. In this

section comparisons by the method of the coefficient of racial likeness* are made between the male Scarborough and seven other British series, viz.:

(1) *Dunstable* (Dingwall & Young, 1933). The skeletons were found as secondary interments in a bell-barrow. The following remarks regarding their age are given in the report on the excavations (Dunning, Wheeler & Dingwall, 1931): 'The absolute dating of either of the two main series of burials (those in trench-graves and those buried superficially) is difficult. So close were they all to the surface of the mound that...it is impossible to affirm that any of the objects found in the same layer were, in the archaeological sense, associated with them....The general tenor of this evidence is that the surface of the mound was disturbed not later than the early Saxon period, and it is inferred that the burials (or the majority of them) are of the fifth or sixth century A.D.' Judging from cranial characters, however, the Dunstable is clearly distinguished from all Anglo-Saxon series.

(2) *Spitalfields* (Morant & Hoadley, 1931). This long series of skulls was recovered when excavations were carried out on the site of Spitalfields Market, London. Their age could not be ascertained as no datable articles were found. The site formed part of a Roman cemetery, and it was almost certainly within the churchyard of St Mary Spittle (1197–1537). It was completely built over by 1688. Comparisons of the measurements show that the Spitalfields has its closest connexions with Pompeian and Etruscan series, while it is little further removed from the following.

(3) *Hythe* (Stoessiger & Morant, 1932). The skulls described form part of the large collection in the ambulatory of St Leonard's Church. The people probably died between A.D. 1100 and 1600. The series was found to have a closer connexion with the Spitalfields than with any other with which it could be compared.

(4) *English Bronze Age* (Morant, 1926). This series is believed to represent the Bronze Age invaders, the Neolithic element having been eliminated by a rough method. Revised means are given in *Biometrika*, **20** B (1928), 368–9.

(5) *Anglo-Saxon* (Morant, 1926). The skulls came from a number of cemeteries.

(6) *British Iron Age* (Morant, 1926). The pooled means represent the south of England better than the north, and England as a whole better than Scotland. The majority of the specimens are of Romano-British date. Revised means are given in *Biometrika*, **20** B (1928), 372–3.

(7) *Whitechapel* (Macdonell, 1904). The series came from a seventeenth-century London plague pit. A few additional measurements of the skulls are given in *Biometrika*, **18** (1926), 28–9.

There are data for other British series, but these seven were selected for comparison with the Scarborough because a preliminary comparison of the mean measurements showed that it diverges clearly from the prevailing type. The Moorfields and Farringdon Street, London, and the Glasgow Scottish series bear a close resemblance to the Whitechapel, and there is good reason to believe that these four represent the racial population which

* With the usual notation, the form of the crude coefficient is

$$\frac{1}{M}\Sigma\left(\frac{n_s n_{s'}}{n_s+n_{s'}}\times\frac{(m_s-m_{s'})^2}{\sigma_s^2}\right)-1\pm 0\cdot 67449\sqrt{\frac{2}{M}}=\frac{1}{M}\Sigma(\alpha)-1\pm 0\cdot 67449\sqrt{\frac{2}{M}}.$$

If $\bar{n}_s$ is the mean number of skulls available for the characters used in the case of the first series, and $\bar{n}_s'$ is the same for the second series, then the 'reduced' coefficient is defined to be

$$50\times\frac{\bar{n}_s+\bar{n}_{s'}}{\bar{n}_s\bar{n}_{s'}}\left\{\frac{1}{M}\Sigma(\alpha)-1\right\}\pm 0\cdot 67449\sqrt{\frac{2}{M}}\times 50\times\frac{\bar{n}_s+\bar{n}_{s'}}{\bar{n}_s\bar{n}_{s'}}.$$

has been spread fairly uniformly over England and the south of Scotland in modern times. The British Iron Age series denotes a population which had very similar characteristics, and the Anglo-Saxon, while standing distinctly apart, was not far removed. The mean male cephalic indices for these six samples have the very restricted range from 74·0 to 75·5. The British Neolithic value is still lower, being 71·7. The Scarborough mean index is 79·0, which is not far removed from the Dunstable (78·7), Spitalfields (79·4), Hythe (82·6), or English Bronze Age (81·3). Racial affinities should not be judged from a single character, but these comparisons are sufficient to show that the Scarborough population cannot have been closely related to that prevailing in England in any period, except possibly the Bronze Age, while it may have been closely connected with the aberrant communities found at Dunstable, London (Spitalfields) and Hythe.

Table 4. *Coefficients of racial likeness between the Scarborough and other male series**

	Scarborough (34·9) and		α's > 12
	Crude C.R.L.	Reduced C.R.L.	
Dunstable (40·2)	0·63 ± 0·19 (26)	1·68 ± 0·51	—
Spitalfields (167·4)	3·68 ± 0·18 (28)	6·37 ± 0·31	L = 20·9, O_2L = 16·8, 100 O_2/O_1, L = 13·1, C = 12·1
Hythe (102·1)	5·18 ± 0·18 (28)	9·96 ± 0·35	L = 42·6, 100 H'/L = 37·7, 100 B/L = 31·6
English Bronze Age (31·1)	3·65 ± 0·19 (25)	10·98 ± 0·57	O_1L = 31·1, 100 O_2/O_1, L = 14·3, B = 12·1
Anglo-Saxon (35·9)	4·32 ± 0·18 (27)	12·26 ± 0·51	100 B/L = 34·6, 100 B/H' = 21·4, B = 15·3, L = 15·3
British Iron Age (55·7)	5·12 ± 0·21 (21)	12·40 ± 0·51	100 B/L = 25·9, L = 21·6, B = 17·5, 100 B/H' = 12·1
Whitechapel (93·3)	7·61 ± 0·19 (25)	14·75 ± 0·37	100 B/L = 56·3, O_1L = 30·2, B = 24·9, L = 15·1
Parisians: l'Ouest (76·7)	1·01 ± 0·24 (16)	2·00 ± 0·48	—
Belgians (66·6)	1·45 ± 0·25 (14)	2·96 ± 0·51	—
Etruscans (79·3)	2·22 ± 0·19 (24)	4·64 ± 0·40	B' = 15·8
S.W. Norwegians (126·0)	2·77 ± 0·19 (25)	5·06 ± 0·35	H' = 18·3, 100 H'/L = 12·3
Finns (120·1)	3·18 ± 0·22 (18)	5·55 ± 0·38	O_2L = 22·8, 100 H'/L = 14·3, L = 13·0
Pompeians (87·1)	3·10 ± 0·21 (20)	5·99 ± 0·41	L = 14·6
Parisians: Cité (57·2)	2·94 ± 0·22 (18)	6·49 ± 0·48	fml = 12·4

* The numbers in brackets following the designations of the series are the average numbers of skulls available for the characters used in computing the coefficients ($\bar{n}$'s). The numbers in brackets following the crude coefficients are the numbers of characters on which they are based. The Farringdon Street standard deviations were used in calculating the coefficients in this table.

Coefficients of racial likeness between the Scarborough and the seven other British series are given in the upper part of Table 4. The crude value with the Dunstable is only just significant (being 3·3 times its probable error), and the reduced value actually indicates closer resemblance than any found between pairs of three series of seventeenth-century London crania. No one of the characters used considered singly shows a difference which is clearly significant. The Scarborough series is seen to be much further removed from both the Spitalfields and Hythe, its divergences from them being appreciably greater than that between the two (reduced coefficient = 4·14). The English Bronze Age gives a rather higher reduced coefficient, and, as was anticipated, still more marked differences are found from the types of the Anglo-Saxon, British Iron Age and Whitechapel series. The distinctions in all these cases depend chiefly on the fact that a few of the characters compared show

markedly significant differences. The Scarborough type differs most clearly from the Spitalfields and Hythe in having a greater calvarial length (associated with marked differences in the height-length and cephalic indices in the latter case), from the English Bronze Age in having a smaller orbital breadth, and from the Anglo-Saxon, British Iron Age and Whitechapel in having a shorter calvarial length but greater breadth, so that its cephalic index is markedly higher than the values for them.

5. Comparisons between the Scarborough and non-British series

Measurements are available for numerous series of skulls representing populations of Western Europe in historical times. When a new series from the region is compared with these it is usually possible to find five or more close connexions, indicated by reduced coefficients of racial likeness less than 5·0. Only one value of this order could be found between the Scarborough and other British series, viz. that of 1·68 with the Dunstable. A rough comparison of the means suggested that it was only likely to give reduced coefficients less than 5·0 with seven European (and non-British) series. These are:

(8) *Parisians: l'Ouest* (Broca, 1873). *Le cimetière de l'Ouest* was opened in 1788 and closed in 1825. The means used are given in *Biometrika*, **23**, 232.

(9) *Belgians* (Heger & Dallemagne, 1881). Measurements are given of four series, viz. three of murderers whose skulls are preserved in the universities of Brussels, Ghent and Liege, respectively, and one of men who died at Brussels. The means of these short series are very similar and they were accordingly pooled, giving the values in Table 3 above. Standard deviations for the total 133 specimens are: $L = 5{\cdot}84 \pm 0{\cdot}24$, $B = 5{\cdot}37 \pm 0{\cdot}22$, $100\ B/L = 3{\cdot}74 \pm 0{\cdot}15$. These are all less than the corresponding constants for the Scarborough series (Table 2). In spite of this the pooling of all the Belgian specimens is not entirely satisfactory, and comparisons with the pooled means are only of provisional value.

(10) *Etruscans* (Schmidt, 1887). The means used are given in *Biometrika*, **20** B, 370.

(11) *South-west Norwegians* (Schreiner, 1939). The means given in Table 3 above were obtained by pooling measurements given for three series from Bergen, Jaeren and Sogn, respectively. These three are very similar in type, judging by the coefficients of racial likeness given by Schreiner, and they differ from the series he describes from south-east Norway in having higher mean cephalic indices and in other respects. The dates of the skulls range from the thirteenth to the nineteenth century.

(12) *Finns* (Morant & Hoadley, 1931, pp. 229, 232). Pooled means were obtained by taking the measurements of several short series given by K. Hällstén and other anthropologists. The majority of the specimens were assigned to the eighteenth and nineteenth centuries.

(13) *Pompeians* (Nicolucci, 1882; Schmidt, 1884). The pooled means used are given in *Biometrika*, **23**, 232.

(14) *Parisians: Cité.* The means used are given in *Biometrika*, **23**, 232. They were obtained from Broca's MS. catalogues. The series represents the population of Paris in the twelfth century.

Reduced coefficients of racial likeness between all pairs of the series 8, 10, 12, 13 and 14 listed above are given by Morant & Hoadley (1931, p. 234). They range from 2·44 to 6·91. Values between them and the Spitalfields series are given in the same table, and their range is from 3·54 to 8·57. The reduced coefficients between the same five series and the Hythe (Stoessiger & Morant, 1932, p. 198) range from 7·61 to 13·52.

Coefficients of racial likeness between the Scarborough and these seven non-British series are given in Table 4. The samples are all fairly adequate in size, but in every case a smaller or larger number of the characters used when possible in calculating the coefficient is not available for them. Hence the values given are only approximations to those which would be obtained for the complete set of characters, but close-approximation is expected when the number is 20 or greater. All the reduced coefficients are seen to be appreciably lower than any found between the Scarborough and a series representing the prevailing population of England in any period. No resemblance with a non-British series is quite as close as that between the Scarborough and Dunstable, but the newly described population appears to have been more closely allied to several continental ones than to the other populations intrusive in Britain, viz. the Spitalfields and Hythe.

6. THE RACIAL RELATIONSHIPS OF THE SCARBOROUGH POPULATION

It is safe to infer from the archaeological evidence that the skulls described in this paper are those of people who died between the early twelfth and mid-sixteenth century. It is not known whether the interments were made at one particular time during this period, or whether they were dispersed over the whole of it. The cemetery was attached to a monastic house which was used up to the time of the Dissolution, and the records suggest that its alien inmates were always few in number. Men predominate but women and children were also buried in the site, and it is probable that the majority of these people were members of the laity of the town.

As far as can be judged from the short series, the men, women and children belonged to the same population. The nature of this group and its racial relationships have to be judged from the forty-three male skulls. This is a small sample for the purpose. Judging from constants for all the characters recorded, there is no appreciable difference between the variation exhibited by the Scarborough series and that of seventeenth-century Londoners who were buried in the graveyard at Farringdon Street. At the same time there is a clear suggestion of racial heterogeneity in the fact that the former has the high standard deviation of 4·41 for the cephalic index, which is known to be the character most likely to reveal exceptional mixture. The Scarborough series is certainly not ideal for comparative purposes, and certain conclusions of a broad nature are all that can be legitimately derived from it.

Comparisons of cranial characters considered singly, and of a number considered in conjunction by using the method of the coefficient of racial likeness, suggest the following conclusions. Considered as a whole and as an English series the Scarborough must be supposed aberrant. The majority of the people represented, if not all of them, cannot have belonged to any of the populations which prevailed in the country at different periods since Mesolithic times. It is possible that a few of them were members of one of those populations, but this cannot be demonstrated. Comparison with the other English series of skulls which represent alien communities shows that the Scarborough bears a close resemblance to the Dunstable (probably of fifth or sixth-century date), but it differs appreciably from both the Spitalfields and Hythe series.

Comparison with continental material shows that the Scarborough bears a close resemblance to several of the mesocephalic populations of Western Europe, a group which embraces French, Belgian and certain Italian series, and to which the inhabitants of *south-west Norway and Finns can be assigned. The closest* connexions—which are of the

same order as that between the Scarborough and Dunstable—are found with a modern Parisian and a modern Belgian series. Several of the characters used when possible in computing coefficients of racial likeness are, however, not available for these two. In view of this fact, and of the defects of the Scarborough series, it would be unwise to lay any stress on the conclusion that it bears a slightly closer resemblance to two particular foreign series than to others belonging to the same group.

A general conclusion which appears to be justified is that the majority of the individuals from Castle Hill described in this paper were descendants of an alien community. It is known from historical evidence that Scarborough was a base for Vikings, who probably occupied the site until the eleventh century. The only known settlement in England of Icelanders was also located there. Hence it is not unlikely that the alien community referred to was pre-Conquest and of Scandinavian origin. Alternatively, it might be suggested that the foreign element was introduced by immigrants from the Low Countries in medieval times. A third possibility is that the peculiarity of the Scarborough population was due to the survival in it of an element derived from the Bronze Age population of Britain, but the evidence does not favour this hypothesis.

The only head measurements of living Yorkshiremen available are those given by Beddoe & Rowe (1907) for a series of ninety subjects from the West Riding. The mean cephalic index, given by Buxton, Trevor & Blackwood (1939), is 78·7. The corresponding cranial index may be supposed to be 77, which is well below the Scarborough value of 79·0, though rather greater than the means for the seventeenth-century London series. It is possible that a population of the Scarborough type had an appreciable effect in determining the characteristics of modern Yorkshiremen.

APPENDIX

Table 5 of individual measurements of the Scarborough skulls

The measurements were taken in accordance with the customary biometric technique. Owing to an error in interpretation, the foraminal breadths (fmb) were inaccurate, and hence they and the indices ($100\, fmb\, fml$) are not recorded. The letters denoting measurements are: C = capacity in cm.^{3} The values given were not found directly but from the reconstruction formulae involving L, B and H' given by Hooke (1926, p. 33). L = maximum glabella-occipital length (Martin 1). B = maximum breadth (M. 8). B' = minimum frontal breadth (M. 9). H' = basio-bregmatic height (M. 17). S_1' = chord nasion to bregma (M. 29). S_2' = chord bregma to lambda (M. 30). S_3' = chord lambda to opisthion (M. 31). S_1 = arc nasion to bregma (M. 26). S_2 = arc bregma to lambda (M. 27). S_3 = arc lambda to opisthion. S = arc nasion to opisthion (M. 25). $\beta Q'$ = transverse arc passing through bregma (M. 24). U = horizontal circumference through the ophryon and above the superciliary ridges (M. 23a). fml = basion to opisthion (M. 7). LB = basion to nasion (M. 5). GL = basion to alveolar point. $G'H$ = nasion to alveolar point (M. 48). GB = facial breadth between lowest points on zygomatic-maxillary sutures (M. 46). J = bizygomatic breadth (M. 45). NH, L = nasion to lowest point on margin of pyriform aperture on left side. NB = maximum breadth of pyriform aperture (M. 54). O_1L = maximum breadth of left orbit (M. 51).

$100\frac{NB}{NH'}, L$	$100\frac{O_2}{O_1}, L$	$100\frac{G_2}{G'_1}$	$N\angle$	$A\angle$	$B\angle$	Remarks
—	—	—	69°	66°	45°	Sagittal suture obliterated. Most teeth lost before death
51·9	74·9	—	64°?	74°·5?	41°·5?	Sagittal and coronal sutures nearly obliterated
52·1?	88·8?	—	—	—	—	Sagittal suture obliterated
—	—	—	69°	62°	49°	Middle aged. Most teeth lost before death
—	—	—	—	—	—	Sagittal, coronal and lambdoid sutures obliterated
53·5?	80·3?	89·9?	—	—	—	Middle aged
—	—	—	57°?	83°?	40°?	Sagittal and lambdoid sutures partly obliterated
52·3?	80·3	105·7?	63°·5	81°	35°·5	Third molars eupting
51·0	73·1?	98·0?	—	—	—	Sagittal and lambdoid sutures obliterated
—	82·1	—	—	—	—	Young adult
51·5	77·8?	91·6	62°	76°	42°	Ageing
48·1	83·9	73·5	70°·5	69°·5	40°	Sagittal, coronal and lambdoid sutures obliterated
56·7	74·7	73·5	69°	79°	32°	Ageing. Edentulous
51·2	69·0	—	63°·5	73°·5	43°	Middle aged
56·2	82·7	—	62°·5	77°·5	40°	Sagittal suture obliterated
49·7	80·3	—	—	—	—	Ageing. Edentulous
45·8	81·4	88·1?	63°·5	74°·5	42°	Sagittal suture nearly obliterated. Lower canines misplaced
48·9?	82·6	93·5	61°·5	72°·5	46°·5	Young adult. Third molars erupting
42·9	93·7	—	61°	71°	48°	Middle aged
—	—	—	62°·5	78°·5	39°	Sagittal and lambdoid sutures obliterated
44·8	88·4?	—	61°·5	75°	43°·5	Middle aged
48·8	77·7	113·4?	63°·5	72°·5	44°	Sagittal, coronal and lambdoid sutures obliterated
—	—	—	—	—	—	Young adult
47·7	—	89·1	61°	82°·5	37°	Middle aged
—	74·0	—	—	—	—	Sagittal suture partly obliterated
—	87·2	93·5	—	—	—	Middle aged
49·7	—	94·1	59°·5	79°·5	41°	Young adult
49·4	98·5	93·8	64°·5	75°	40°·5	Sagittal suture closed
44·4	81·3	88·2	65°·5	75°·5	39°	Young adult
50·5	78·0	105·0	59°	74°·5	46°·5	Ageing
—	90·2?	—	—	—	—	Middle aged
47·3?	96·1	—	60°·5	72°·5	47°	Sagittal and lambdoid sutures closing
45·1	80·1	—	60°?	81°·5?	38°·5?	Ageing
45·1	80·8	—	60°	75°	45°	Sagittal suture nearly closed
50·5	78·3	90·3	68°	69°·5	42°	Middle aged
46·5	87·6	—	61°·5?	78°?	40°·5?	Ageing
47·3	83·6?	72·9?	66°·5	71°	42°·5	Sagittal suture partly closed
48·4	79·8?	—	—	—	—	Aged. Edentulous
52·2	82·8	—	—	—	—	—
51·6	84·8	85·7	70°·5	70°·5	39°	Ageing
49·8	76·1	—	65°·5	71°·5	43°	Middle aged
—	76·7	—	—	—	—	Ageing. Edentulous
46·8	61·7	—	60°	77°·5	42°·5	Middle aged
47·5	88·4	—	65°	71°	44°	Sagittal and lambdoid sutures nearly obliterated
51·8	80·7	73·2	68°·5	71°·5	40°	Middle aged
47·3	85·3	87·0	62°·5	70°·5	47°	—
52·8	81·1	110·6	65°·5	79°·5	35°	Young adult
—	—	—	—	—	—	Middle aged
44·1	90·0	79·6	68°?	69°?	43°?	Middle aged
—	—	—	—	—	—	Sagittal and lambdoid sutures closed
50·6	82·0?	—	67°·5	75°	37°·5	Ageing
—	94·8	—	—	—	—	Middle aged
49·8	97·5?	86·3?	65°	73°	42°	Young adult. Posthumously deformed
44·4	78·1	59·1	65°·5	73°	41°·5	Middle aged
—	84·1	—	—	—	—	Middle aged
54·1?	86·8	100·0	59°·5?	74°·5?	46°?	Young adult. Metopic
52·6	90·7	—	60°?	82°·5?	37°·5?	Middle aged
52·1	100·7	57·0	72°·5	68°·5	39°	—
46·5	87·5	—	61°?	78°?	41°?	Ageing
47·0	88·9	—	57°	77°·5	45°	Middle aged

O_2L = maximum height of left orbit (M. 52). G_1' = length of palate from orale to staphylion (M. 62). $Oc.I.$ = occipital index which is

$$100 \frac{S_3}{S_3'} \sqrt{\left(\frac{S_3}{24\,(S_3 - S_3')}\right)}.$$

Values were found with the aid of a table of the function given in *Biometrika*, **13**, 261. $N\angle$, $A\angle$ and $B\angle$ are the angles of the triangle of which the nasion, alveolar point and basion are the apices. They were found from the sides $G'H$, GL and LB with the aid of a trigonometer.

The mean cephalic index for seven juvenile skulls was found to be 80·7.

REFERENCES

BEDDOE, J. & ROWE, J. H. (1907). The ethnology of West Yorkshire. *Yorks. Archaeol. J.* **19**.

BROCA, P. (1873). La race celtique ancienne et moderne. *Rev. Anthrop.* **2**, 577–628.

BUXTON, L. H. D., TREVOR, J. C. & BLACKWOOD, BEATRICE (1939). Measurements of Oxfordshire villagers. *J. R. Anthrop. Inst.* **69**, 1–10.

DINGWALL, DORIS & YOUNG, M. (1933). The skulls from excavations at Dunstable, Bedfordshire. *Biometrika*, **25**, 147–57.

DUNNING, G. C., WHEELER, R. E. M. & DINGWALL, DORIS (1931). A barrow at Dunstable, Bedfordshire. *Archaeol. J.* **88**, 193–217.

HEGER, M. P. & DALLEMAGNE, M. G. (1881). Étude sur les caractères craniologiques d'une série d'assassins exécutés en Belgique. *Ann. Univ. Bruxelles.*

HOOKE, BEATRIX G. E. (1926). A third study of the English skull with special reference to the Farringdon Street crania. *Biometrika*, **18**, 1–55.

MACDONELL, W. R. (1904). A study of the variation and correlation of the human skull, with special reference to English crania. *Biometrika*, **3**, 191–244.

MORANT, G. M. (1926). A first study of the craniology of England and Scotland from neolithic to early historic times, with special reference to the Anglo-Saxon skulls in London museums. *Biometrika*, **18**, 56–98.

MORANT, G. M. & HOADLEY, M. F. (1931). A study of the recently excavated Spitalfields crania. *Biometrika*, **23**, 191–248.

NICOLUCCI, G. (1882). Crania pompeiana. *Atti R. Acad. Sc. Fis. Mat. Napoli*, **9**.

ROWNTREE, A. (1931). *The History of Scarborough.*

SCHMIDT, E. (1884). Die antiken Schädel Pompejis. *Arch. Anthrop.* **15**, 229–57.

—— (1887). *Die anthropologischen Sammlungen Deutschlands* (Leipzig Catalogue).

SCHREINER, K. E. (1939). Crania Norvegica. I. *Inst. Samm. Kulturforskning*, Serie B, *Skrifter*, **36**, 1–201.

STOESSIGER, BRENDA N. & MORANT, G. M. (1932). A study of the crania in the vaulted ambulatory of Saint Leonard's Church, Hythe. *Biometrika*, **24**, 135–202.

TILDESLEY, M. L. (1928). Skull with post-mortem trepan, assigned to the mid-sixteenth century. *Proc. R. Soc. Med., Sec. Hist. Med.* **21**, 62–5.

A STUDY OF THE CHINESE HUMERUS

By T. L. WOO, Ph.D.

Fellow of the Institute of History and Philology, Academia Sinica

1. *Introduction.* In general the published anthropometric data relating to the humerus are far less extensive than those for the cranium, and this is particularly so in the case of Chinese material. Several good series of Chinese crania have been described (see Woo & Morant, 1932), but there are few memoirs which deal with series of Chinese skeletons in any detail. The material treated in Black's study (1925) comprises two short series from late prehistoric sites and one representing the modern inhabitants of the north China plain: all three relate to the northern part of the country. I have been able to study two series of skeletons from central China. One of these is short and of medieval date. The other is modern, and it is long enough to serve most statistical purposes. The present paper provides a description of the humeri, and it is hoped that detailed studies of the other bones of the skeletons will be published later. The modern material was extensive enough to make possible a more thorough statistical examination of bilateral and sexual differences than any which appears to have been provided previously for the humerus. Comparisons with other racial series were restricted by the fact that the writer had limited access to the literature while working under difficult war conditions.

2. *The new material.* Two series of Chinese humeri preserved in the Museum of the Institute of History and Philology, Academia Sinica, are described in this paper. The writer is greatly indebted to the authorities of the Institute, and particularly to Professors Fu Ssûnien and Li Chi, for granting him permission to study these bones. The adult specimens only were examined and they were sexed anatomically, with the aid of the crania and other parts of the skeletons. Most of the humeri are well preserved, but one or both extremities are defective in a few cases. The series came from:

(i) Hsiao T'un, Anyang. Skeletal remains representing the Sui-T'ang dynasties (A.D. 581–899) were excavated from 1929 to 1932 by Dr Li Chi and other Fellows of the Archaeological Section of the Institute. The specimens came from several cemeteries, and an account of the excavations will be given in a report on the crania which the writer is preparing. There are eleven pairs of male and seven pairs of female humeri. Measurements of unpaired bones in the two series were taken but they are not used in this paper.

(ii) Hsiu Chiu Shan, south of Hsia Kuan, Nanking. The skeletons are modern and they were obtained by the writer in 1936 from unclaimed graves used by the poorer classes. It is known that some of the people came from the eastern part of China and some are from unknown localities. There are seventy pairs of male and forty-three pairs of female humeri.

3. *The measurements recorded.* Martin's technique (1928) was followed and a selection of the measurements which appear to be most useful for comparative purpose was made. They are:

(1) Maximum length from the highest point on the head to the lowest point on the trochlea, measured with the osteometric board (M. 1).

(2) Total length from the highest point on the head to the lowest point on the capitulum, taken with the axis of the shaft parallel to the side wall of the board (M. 2).

(3) Breadth of the proximal epiphysis from the most medial point on the articular surface of the head to the most lateral point on the greater tuberosity, taken on the board with the shaft of the bone parallel to the side wall (M. 3).

(4) Breadth of the distal epiphysis, i.e. the maximum breadth between the condyles taken on the board with the bone in the same position as for (3) (M. 4).

(5) Maximum diameter at the middle of the shaft (determined from the maximum length (1)) taken in any direction with small callipers (M. 5).

(6) Minimum diameter at the middle of the shaft taken without reference to the direction of the maximum diameter with small callipers (M. 6).

(7) Circumference at the middle of the shaft measured with a steel tape (M. 7*a*).

(8) Minimum circumference measured with a steel tape and usually found at the second third, distal to the deltoid eminence (M. 7).

(9) Maximum sagittal diameter of the head from the highest point on the margin of the articular surface of the head to the lowest point on the same margin, taken with small callipers in a plane parallel to the long axis of the bone (M. 10).

(10) Maximum transverse diameter of the head perpendicular to (9) (M. 9).

(11) Circumference of the head measured round the margin of the articular surface with a slip of paper (M. 8).

(12) Angle of torsion. With the bone in the standard horizontal position, this is the angle between lines representing the axes of the articular surfaces of the proximal and distal ends projected on to a

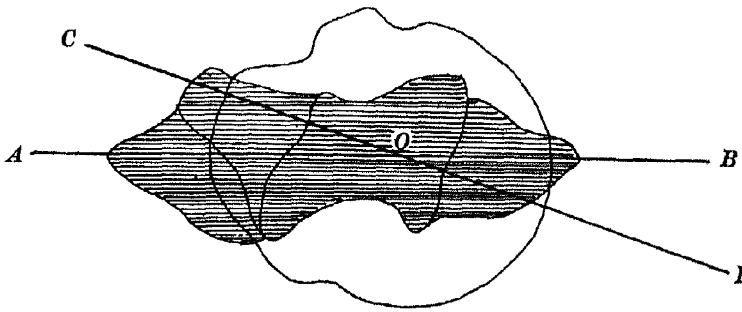


Fig. 1. Contours of the two ends of a humerus superposed, showing the angle of torsion ($< AOD$) between their axes. (*After Martin.*)

vertical plane. Fig. 1 illustrates the way in which it is measured. The line CD is the axis determined by the 'centre' of the head and the 'mid-point' of the greater tuberosity, so that it appears to divide the outline of the proximal extremity into two equivalent areas. The line AB is the axis of the distal extremity determined so that it appears to halve the surface of the trochlea and capitulum which is seen. The two lines intersect at O and the angle AOD is defined to be the angle of torsion (M. 18). Some authors use the supplement of this angle instead. In practice the bone is held vertical by a clamp, and two fine knitting needles are used to represent the axes. Their directions are marked on a sheet of paper with the aid of a parallelograph and the angle between them is measured with a protractor (M. 18).

Three indices derived from pairs of the direct measurements defined above were used, viz.:

(13) Cross-section index of the diaphysis $= 100 \times 6/5$.

(14) Caliber index $= 100 \times$ minimum circumference (8)/maximum length (1).

(15) Head index $= 100 \times 10/9$.

The measurements taken on the board and the circumferences were recorded to the nearest 0·5 mm., and the chords, found with small callipers provided with a vernier scale, were recorded to the nearest 0·1 mm. Constants derived from the measurements are given in Table 1.

4. *Bilateral comparisons.* For the eleven measurements of size the mean on the right side exceeds the corresponding value on the left in the case of both male and female series with only one exception, viz. the minimum diameter at the middle of the shaft for the female Hsiu Chiu Shan series. It had been found from material representing other races that on the average the right humerus is appreciably larger than the left in nearly all respects.

Table 1. *Constants (with probable errors) for two Chinese series of humeri: paired bones only*

Constant	Means							
Sex	Male				Female			
Series	Hsiao T'un		Hsiu Chiu Shan		Hsiao T'un		Hsiu Chiu Shan	
Side	R.	L.	R.	L.	R.	L.	R.	L.
1. Maximum length	312·9 (7)	308·9 (7)	310·32 ± 1·25 (61)	306·72 ± 1·20 (61)	286·4 (5)	282·6 (5)	284·94 ± 1·81 (38)	281·71 ± 1·74 (38)
2. Total length	307·0 (7)	304·0 (7)	304·50 ± 1·20 (60)	301·88 ± 1·16 (60)	278·9 (5)	276·5 (5)	278·45 ± 1·76 (36)	276·16 ± 1·68 (36)
3. Br. prox. epiphysis	48·5 (6)	47·9 (6)	48·03 ± 0·19 (59)	47·81 ± 0·20 (59)	43·8 (5)	43·3 (5)	43·53 ± 0·32 (36)	43·18 ± 0·29 (36)
4. Br. dist. epiphysis	59·8 (6)	59·2 (6)	59·40 ± 0·23 (64)	58·79 ± 0·22 (64)	54·3 (7)	53·5 (7)	53·98 ± 0·44 (38)	53·24 ± 0·42 (38)
5. Max. diam. mid. shaft	23·4 (11)	22·6 (11)	22·37 ± 0·09 (61)	21·65 ± 0·09 (61)	20·0 (7)	19·0 (7)	19·97 ± 0·16 (38)	18·99 ± 0·15 (38)
6. Min. diam. mid. shaft	18·1 (11)	17·6 (11)	17·15 ± 0·11 (61)	17·01 ± 0·10 (61)	13·9 (7)	13·5 (7)	13·62 ± 0·12 (38)	13·65 ± 0·11 (38)
7. Circum. mid. shaft	68·5 (11)	66·9 (11)	67·81 ± 0·43 (61)	66·24 ± 0·45 (61)	58·5 (7)	56·9 (7)	57·41 ± 0·46 (38)	55·93 ± 0·46 (38)
8. Min. circum. shaft	64·8 (11)	63·1 (11)	63·98 ± 0·38 (70)	62·58 ± 0·35 (70)	56·5 (7)	55·0 (7)	56·04 ± 0·37 (42)	54·53 ± 0·38 (42)
9. Sag. diam. head	46·0 (7)	44·0 (7)	46·08 ± 0·18 (65)	44·09 ± 0·18 (65)	39·6 (6)	38·4 (6)	39·33 ± 0·21 (43)	38·21 ± 0·22 (43)
10. Trans. diam. head	43·8 (7)	42·8 (7)	43·64 ± 0·19 (64)	42·70 ± 0·17 (64)	36·5 (6)	36·0 (6)	36·31 ± 0·19 (43)	35·82 ± 0·19 (43)
11. Circum. head	138·2 (7)	136·5 (7)	137·56 ± 0·53 (65)	135·96 ± 0·53 (65)	123·0 (6)	122·1 (6)	121·69 ± 0·77 (43)	120·82 ± 0·75 (43)
12. Angle of torsion	149°·3 (6)	148°·6 (6)	149°·30 ± 0·73 (61)	148°·81 ± 0·70 (61)	146°·2 (7)	145°·6 (7)	146°·95 ± 0·84 (38)	146°·17 ± 0·71 (38)
13. Index of diaphysis	77·1 (11)	78·2 (11)	76·66 ± 0·29 (61)	78·54 ± 0·28 (61)	68·9 (7)	70·8 (7)	68·28 ± 0·25 (38)	71·62 ± 0·30 (38)
14. Caliber index	20·8 (7)	20·4 (7)	20·72 ± 0·15 (61)	20·34 ± 0·15 (61)	19·8 (5)	19·3 (5)	19·70 ± 0·14 (38)	19·30 ± 0·15 (38)
15. Head index	95·6 (7)	96·7 (7)	95·33 ± 0·13 (64)	96·29 ± 0·13 (64)	92·0 (6)	93·8 (6)	92·18 ± 0·24 (43)	93·84 ± 0·20 (43)

Table 1 (*continued*)

Constant	Standard deviations				Bilateral correlations	
Sex	Male		Female			
Series	Hsiu Chiu Shan					
Side	R.	L.	R.	L.	Male	Female
1. Maximum length	14·49 ± 0·88	13·92 ± 0·85	16·53 ± 1·28	15·87 ± 1·23	0·96 ± 0·01	0·97 ± 0·01
2. Total length	13·80 ± 0·85	13·32 ± 0·82	15·67 ± 1·24	14·97 ± 1·19	0·96 ± 0·01	0·97 ± 0·01
3. Br. prox. epiphysis	2·13 ± 0·13	2·23 ± 0·14	2·82 ± 0·22	2·58 ± 0·21	0·96 ± 0·01	0·98 ± 0·01
4. Br. dist. epiphysis	2·67 ± 0·16	2·60 ± 0·16	3·99 ± 0·31	3·88 ± 0·30	0·94 ± 0·01	0·91 ± 0·02
5. Max. diam. mid. shaft	1·09 ± 0·07	1·09 ± 0·07	1·42 ± 0·11	1·35 ± 0·10	0·83 ± 0·03	0·78 ± 0·04
6. Min. diam. mid. shaft	1·24 ± 0·08	1·10 ± 0·07	1·07 ± 0·08	1·02 ± 0·08	0·51 ± 0·06	0·50 ± 0·08
7. Circum. mid. shaft	4·97 ± 0·30	5·20 ± 0·32	4·19 ± 0·32	4·24 ± 0·33	0·54 ± 0·06	0·52 ± 0·08
8. Min. circum. shaft	4·76 ± 0·27	4·38 ± 0·25	3·54 ± 0·26	3·67 ± 0·27	0·56 ± 0·06	0·56 ± 0·07
9. Sag. diam. head	2·18 ± 0·13	2·12 ± 0·13	2·06 ± 0·15	2·18 ± 0·16	0·89 ± 0·02	0·76 ± 0·04
10. Trans. diam. head	2·20 ± 0·13	2·02 ± 0·12	1·86 ± 0·14	1·88 ± 0·14	0·86 ± 0·02	0·85 ± 0·03
11. Circum. head	6·34 ± 0·38	6·38 ± 0·38	7·44 ± 0·54	7·05 ± 0·51	0·96 ± 0·01	0·96 ± 0·01
12. Angle of torsion	8°·45 ± 0·52	8°·05 ± 0·49	7°·65 ± 0·59	6°·45 ± 0·50	0·46 ± 0·07	0·46 ± 0·09
13. Index of diaphysis	3·40 ± 0·21	3·23 ± 0·20	2·26 ± 0·17	2·71 ± 0·21	0·46 ± 0·07	0·54 ± 0·08
14. Caliber index	1·77 ± 0·11	1·73 ± 0·11	1·26 ± 0·10	1·39 ± 0·10	0·66 ± 0·05	0·60 ± 0·07
15. Head index	1·57 ± 0·09	1·57 ± 0·09	2·31 ± 0·17	1·98 ± 0·14	0·42 ± 0·07	0·55 ± 0·07

Consistent bilateral differences between means are also found in the case of the four measurements of shape. For all four series the angle of torsion and caliber index have greater values on the right, and the diaphysial and head indices have greater values on the left. Differences having the same signs have generally been found for other series.

Standard deviations and bilateral correlations are given in Table 1 for the bones from Hsiu Chiu Shan. With the aid of these constants the statistical significance of the bilateral differences between the means can be tested.* The differences divided by their probable errors give the following values for the series in question, a negative sign indicating that the left mean is greater than the right:

No. of measurement	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
♂	10·6	7·7	3·7	7·6	14·4	1·4	3·7	4·1	24·9	9·4	11·4	0·7	−6·3	3·2	−6·9
♀	7·3	5·3	3·5	7·4	8·2	−0·3	3·4	4·4	7·5	4·9	4·0	1·0	−12·4	3·1	−7·9

The corresponding values of the ratios indicate a close agreement between the two sexes, and the majority provide clear evidence of asymmetry for both. It may be said, without regard to sex or race, that on the average the right humerus tends to be larger than the left in all respects, but that the side difference is much less marked for some transverse measurements of the shaft (nos. 6, 7 and 8), other than the maximum diameter of the section at the middle (5), than for the length of the bone and most measurements of its extremities. The torsion (12) tends to be greater on the right than on the left, and the same is true for the caliber index, or index of robustness (14). By far the most significant bilateral differences in shape, however, are for the index (13) expressing the minimum diameter at the middle of the shaft as a percentage of the maximum, which is greater on the left, and for the index (15) expressing the transverse as a percentage of the sagittal diameter of the head, which shows a difference of the same sign.

The significance of the bilateral differences between the standard deviations given in Table 1 for the longer of the series was estimated, as in the case of the means, by dividing each difference by its probable error.† For the fifteen characters, the highest of these ratios found for the male comparisons is 2·0, so all the differences must be considered quite insignificant. The highest ratio for the female comparisons is 4·0 (breadth of proximal epiphysis) and the next highest is 2·0. The former might be considered to indicate a significant difference if it were found by itself, but no importance can be attached to it considered as the extreme in a set of ratios. So far there appears to be no evidence of distinction between the variabilities of characters of the humerus on the right and left sides. For the eleven measurements of size, however, the standard deviations for the right side tend to be greater than those for the left. For these male constants the right value is the greater in seven cases, there is equality for one and the left is the greater in three cases: for the female constants the right is the greater in seven cases and the left is the greater in the remaining four. The tendency for the absolute variabilities of size measurements to be greater on the right accords with the fact that the right humerus is larger than the left, and

* The formula for the probable error of a difference is $0{\cdot}6745\sqrt{(\sigma_R^2+\sigma_L^2-2r_{RL}\sigma_R\sigma_L)}/\sqrt{n}$.

† The formula used for the probable error of the difference is $0{\cdot}6745\sqrt{(\sigma_R^2+\sigma_L^2-2r_{RL}^2\sigma_R\sigma_L)}/\sqrt{(2n)}$.

measures of relative variability (coefficients of variation) would doubtless make less distinction between the sides. A few markedly significant bilateral differences have been found for the absolute and relative variabilities of single bones of the cranium (Woo, 1931), but these criteria would not be expected to indicate distinction for series as short as those of the humeri from Hsiu Chiu Shan.

The bilateral correlations are given in the last two columns of Table 1. The average values for the male constants are 0·85 for the eleven size measurements and 0·50 for the four indices, the corresponding female means being 0·80 and 0·54. These may be compared with the means of 0·82 found for twenty-five absolute measurements of single bones of the cranium and 0·74 for twelve indices derived from them in the case of a long Egyptian series (Pearson & Woo, 1935). The correlations are high for the lengths, for all the absolute measurements of the epiphyses and for the maximum diameter at the middle of the shaft, but low—considered as bilateral correlations—for the minimum diameter of the same section and for the two circumferences of the shaft.

5. *Sexual comparisons.* All the male means in Table 1 are clearly greater than the corresponding female values. The following sex ratios for the absolute measurements (male mean/female mean) are found for the Hsiu Chiu Shan series of right bones:

No. of measurement	1	2	3	4	5	6	7	8	9	10	11
Sex ratio	1·089	1·094	1·103	1·100	1·120	1·259	1·181	1·142	1·172	1·202	1·130

The ratios are lower for the lengths (1 and 2) than for the other characters, but all are decidedly high for skeletal measurements. The average for the eleven characters is 1·145, and the average sex ratio for fifty-nine absolute measurements of the cranium given by the unpublished series from Hsiao T'un is 1·052. For all the series in the table the male means of the angle of torsion and of the three indices are greater than the corresponding female values, and for the Hsiu Chiu Shan all these differences are seen to be markedly significant except the two between the angles of torsion. Judging from the new Chinese series, the type of the male humerus has less torsion than the female, a less flattened section of the shaft, with greater girth relative to the length of the bone, and a head which approaches more nearly a circular form. Some of these relations do not accord with those found for other series. Table 3 gives the means of the Hsiu Chiu Shan and those of longer series of Lapp (Schreiner, 1931) and Norwegian (Wagner, 1927) bones. The three agree in showing a greater male than female mean for the caliber index, and all the differences may be supposed markedly significant. The index of the diaphysis is markedly greater for the Chinese male than for the female series; the Norwegian show a difference of the same sign which is much smaller but still probably significant, and the male and female means for Lapps are almost identical. Data for other races suggest that the diaphysial index is normally greater for males than for females, on the average. For the Chinese series there is also clear male dominance in the case of the head index, but the differences for Lapps and Norwegians are quite insignificant. Even less consistency is found for the angle of torsion. The male mean for right humeri of the Hsiu Chiu Shan series exceeds the female by 2°·4, which is 2·2 times the probable error of the difference, but for both Lapps (Δ/P.E. $\Delta = 3{\cdot}3$) and Norwegians (5·4) the female mean exceeds the male. It is possible that there are racial

distinctions in the sexual difference of humeral characters, but more extensive material would be required to establish this point.

No significant differences are found for the Hsiu Chiu Shan series between the corresponding male and female standard deviations, or between the bilateral correlations.

6. *Racial comparisons of mean measurements.* Metrical data relating to series of bones are far less extensive for the humerus than for the cranium, and considerably less numerous for the humerus than for the femur. The value of the available material is lessened appreciably owing to the fact that different definitions of measurements have been used by different workers, and also because constants are only given for sides, or sexes, or both, combined in the case of some of the series. As there are significant side and sexual differences, it is most undesirable that any such combinations should be made.

Table 2 gives mean measurements for the two Chinese series measured by Black (1925) and the two described for the first time in the present paper.* Black gives data for two other characters, but the definitions used do not accord with Martin's. Close agreement between the means is found in nearly all cases in spite of the fact that three of the series are very small. There is no suggestion of a significant difference for any one of the five characters.

In Table 3 comparison is made between the modern series from Nanking, one of Lapps described by Schreiner (1931) and one of Norwegians described by Wagner (1927). All the measurements recorded are involved in this case and all the series are adequate in length for both sexes. No significant differences are found between the lengths of the Chinese and Lapp series. In the case of the male means of the absolute measurements differences exceeding 3·5 times their probable errors are only found for the maximum diameter of the shaft at the middle (Δ/P.E. $\Delta = 6{\cdot}0$), for the minimum diameter of the shaft (4·7), and for the transverse diameter of the head (4·7). For the first two of these measurements the Lapp mean exceeds the Chinese, and for the third the Chinese is the greater. In the case of the female means of the absolute measurements the significant differences are for the minimum diameter of the shaft at the middle (15·1), the circumference of the shaft (4·8), the sagittal diameter of the head (3·9), and the transverse diameter of the head (5·2), the Lapp mean being the greater for all four. These relations are appreciably different for the two sexes and the measurements of shape show the same discordance. For both males (7·6) and females (11·7) the angle of torsion is greater for the Lapp bones. For males the difference for the index of the diaphysis is insignificant, and for females the Lapp mean exceeds the Chinese by an amount which is 23·0 times its probable error. The head index also show unexpected relations, the Chinese male mean being significantly greater than the Lapp (8·3), while the Lapp female mean is significantly greater than the Chinese (5·6).

Generalization is difficult in view of these results. One clear conclusion, however, is that the average types of the humerus for two racial populations may differ with marked significance in the case of some features when the total lengths are undifferentiated. Judging from the data for both sexes, it can be stated that the Lapp bone tends to have a more massive section of the mid-shaft and a greater angle of torsion than the Chinese, but no other clearly significant and consistent differences are found. In the comparison of the two series, there is marked disagreement between the male and female relations in the case of the index of the diaphysis and of the head index. It can be seen from Table 3 that the

* Measurements of other series of Chinese humeri have probably been given by Haberer (1892) and Kurz (1922), but I was unable to consult these sources.

Table 2. *Mean measurements of Chinese series of right humeri**

		1. Maximum length	4. Br. dist. epiphysis†	5. Max. diam. mid. shaft	6. Min. diam. mid. shaft	13. Index of diaphysis (6/5)
♂	Yang Shao: prehistoric	332·0 (2)	61·2 (4)	23·5 (4)	17·7 (4)	75·4 (4)
	Hsiao T'un: medieval	312·9 (7)	59·8 (6)	23·4 (11)	18·1 (11)	77·1 (11)
	North China: modern	311·7‡ (20)	60·2 (20)	22·7 (20)	17·2 (20)	75·7 (20)
	Hsiu Chiu Shan: modern	310·3 ± 1·2 (61)	59·4 ± 0·23 (64)	22·4 ± 0·09 (61)	17·1 ± 0·11 (61)	76·7 ± 0·29 (61)
♀	Yang Shao: prehistoric	284·2 (4)	55·5 (4)	20·4 (5)	14·4 (5)	70·7 (5)
	Hsiao T'un: medieval	286·4 (5)	54·3 (7)	20·0 (7)	13·9 (7)	68·9 (7)
	North China: modern	285·6 (3)	50·3 (3)	19·0 (3)	13·0 (3)	68·5 (3)
	Hsiu Chiu Shan: modern	284·9 ± 1·8 (38)	54·0 ± 0·44 (38)	20·0 ± 0·16 (38)	13·6 ± 0·12 (38)	68·3 ± 0·25 (38)

* The Yang Shao Tsun and north China series are Black's (1925) and the other two are described in the present paper.

† Black determined the maximum breadth of the distal epiphysis with small callipers and I followed Martin in measuring it on the board, but readings obtained in the two ways may be supposed fairly comparable.

‡ Stevenson (1929) gives a maximum length of 310·7 ± 1·2 for forty-eight male and right humeri in the collection of modern Chinese skeletons at the Peiping Union Medical College. This series almost certainly includes the series measured earlier by Black.

Table 3. *Mean measurements (with probable errors) of the modern Chinese and two other racial series of humeri*

Characters*	Male			Female		
	Chinese (Hsiu Chiu Shan)	Lapp	Norwegian	Chinese (Hsiu Chiu Shan)	Lapp	Norwegian
1. Maximum length (1)	310·3 ± 1·25 (61)	308·0 ± 0·85 (148)	337·3 ± 0·86 (162)	284·9 ± 1·81 (38)	284·8 ± 0·81 (131)	312·7 ± 0·97 (158)
2. Total length (2)	304·5 ± 1·20 (60)	302·5 ± 0·86 (146)	331·0 (157)	278·5 ± 1·76 (36)	280·5 ± 0·80 (127)	307·6 (148)
3. Br. prox. epiphysis (3)	48·0 ± 0·19 (59)	48·1 ± 0·20 (97)	51·9 ± 0·15 (147)	43·5 ± 0·32 (36)	43·3 ± 0·15 (99)	46·4 ± 0·16 (142)
4. Br. dist. epiphysis (4)	59·4 ± 0·23 (64)	60·5 ± 0·22 (103)	65·1 ± 0·19 (152)	54·0 ± 0·44 (38)	54·2 ± 0·20 (82)	57·7 ± 0·22 (149)
5. Max. diam. mid. shaft (5)	22·4 ± 0·09 (61)	23·2 ± 0·10 (141)	24·0 ± 0·08 (161)	20·0 ± 0·16 (38)	20·4 ± 0·08 (120)	21·1 ± 0·09 (158)
6. Min. diam. mid. shaft (6)	17·1 ± 0·11 (61)	17·8 ± 0·10 (141)	19·0 ± 0·07 (161)	13·6 ± 0·12 (38)	15·7 ± 0·07 (120)	16·2 ± 0·08 (158)
7. Circum. mid. shaft (7a)	67·8 ± 0·43 (61)	68·1 ± 0·28 (140)	72·0 ± 0·21 (160)	57·4 ± 0·46 (38)	59·9 ± 0·24 (120)	62·1 ± 0·23 (158)
8. Min. circum. shaft (7)	64·0 ± 0·38 (70)	64·4 ± 0·26 (141)	67·6 ± 0·20 (162)	56·0 ± 0·37 (42)	56·6 ± 0·22 (120)	57·9 ± 0·21 (158)
9. Sag. diam. head (10)	46·1 ± 0·18 (65)	45·5 ± 0·18 (115)	48·5 (151)	39·3 ± 0·21 (43)	40·3 ± 0·15 (96)	43·2 (145)
10. Trans. diam. head (9)	43·6 ± 0·19 (64)	42·3 ± 0·20 (94)	45·2 (135)	36·3 ± 0·19 (43)	37·6 ± 0·16 (84)	40·4 (121)
11. Circum. head (8)	137·6 ± 0·53 (65)	138·0 ± 0·64 (89)	147·7 ± 0·45 (116)	121·7 ± 0·77 (43)	122·8 ± 0·47 (79)	131·1 ± 0·55 (107)
12. Angle of torsion (18)	149°·3 ± 0·73 (61)	156°·1 ± 0·52 (125)	161°·1 ± 0·45 (149)	146°·9 ± 0·84 (38)	158°·6 ± 0·54 (113)	165°·0 ± 0·56 (140)
13. Index of diaphysis (6/5)	76·7 ± 0·29 (61)	77·0 ± 0·32 (141)	78·8 (161)	68·3 ± 0·25 (38)	77·1 ± 0·29 (120)	77·0 (158)
14. Caliber index (7/1)	20·7 ± 0·15 (61)	20·9 ± 0·08 (140)	20·1 (162)	19·7 ± 0·14 (38)	19·9 ± 0·08 (119)	18·5 (158)
15. Head index (9/10)	95·3 ± 0·13 (64)	93·1 ± 0·23 (93)	93·4 (135)	92·2 ± 0·24 (43)	94·2 ± 0·26 (84)	93·6 (121)

* The numbers preceding the measurements are those used in this paper, and the numbers in brackets following the measurements are those of Martin's list.

corresponding male and female means of these two characters differ insignificantly in the case of the Lapp series. In the case of the Norwegian, the sexual difference is insignificant for the head index and probably significant for the index of the diaphysis. In the case of the Chinese the differences are far larger and both are markedly significant. Two possible explanations of this situation may be suggested. The first is that the male and female Chinese series do not represent precisely the same population. The second is that the two measures of shape referred to may be influenced to some extent by conditions of life which were much more dissimilar for men and women in the Chinese than in the Lapp and Norwegian populations. The former appears to be the more plausible hypothesis, but more extensive data for racial series of humeri will be needed to clarify the position.

Probable errors are not provided for some of the Norwegian means quoted in Table 3, but it is clear that they exceed the corresponding Chinese and Lapp values with marked significance in the case of all the absolute measurements. The Norwegian bone is the largest in all respects. It also has the largest angle of torsion in the case of both sexes. Less significant differences are found for the three indices, but they still differentiate the Norwegian from both Chinese and Lapp means in one or more instances. Racial types of the humerus are evidently distinguished in shape as well as in size.

Table 4. *Frequencies of septal apertures for adult series of humeri (R. and L.)*

		Reported by	Total no.		No. with apertures		%	
			♂	♀	♂	♀	♂	♀
1	Prehistoric Chinese	Black	7	13	1	1	14·3 ± 8·9	7·7 ± 5·0
2	Medieval Chinese	Woo	21	15	2	2	9·5 ± 4·3	13·3 ± 5·9
3	Modern Chinese	Woo	162	94	14	10	8·6 ± 1·5	10·6 ± 2·1
4	Modern Chinese	P'an Ming-tzu	226	14	17	4	7·5 ± 1·2	28·6 ± 8·1
1–4	All Chinese	Black, etc.	416	136	34	17	8·2 ± 0·9	12·5 ± 1·9
5	Koreans	Akabori	233	20	16	3	6·9 ± 1·1	15·0 ± 5·4
6	Japanese	Akabori	462	250	45	69	9·7 ± 0·9	27·6 ± 1·9
7	Ainu	Akabori	108	74	15	14	13·9 ± 2·2	18·9 ± 3·0
8	'Whites' in U.S.A.	Hrdlička & Trotter	3213	1095	133	102	4·1 ± 0·2	9·3 ± 0·6
1–7	Orientals	Akabori, etc.	1219	480	110	103	9·0 ± 0·6	21·5 ± 1·3
9	American Negroes	Hrdlička & Trotter	704	252	67	66	9·5 ± 0·7	26·2 ± 1·9
10	Eskimos	Hrdlička	588	569	56	171	9·5 ± 0·8	30·1 ± 1·3
11	American Indians	Hrdlička	1665	1420	308	604	18·5 ± 0·6	42·5 ± 0·9
12	Ancient Egyptians	Hrdlička	264	300	87	176	33·0 ± 2·0	57·0 ± 1·9

7. *Anomalies*. The only anomaly of the humerus of anthropological interest for which extensive records are available is the opening in the bony septum that separates the coronoid from the olecranon fossa, known as a septal aperture. In a comprehensive memoir Hrdlička (1932) has collected data from the earlier literature relating to the frequency of occurrence of this condition in racial series of bones and added extensive observations of his own. He concludes that the anomaly is inherited and his figures show that it makes suggestive racial distinctions. For the same population the aperture occurs more frequently in left bones than in right, and also more frequently in females than in males. In view of these differences it is most desirable that data should be given for sides separately and for

sexes separately. Unfortunately the only records now available for the new Chinese series relate to right and left bones taken together for each sex considered separately.*

The figures are given in Table 4, together with others provided by Black (1925), P'an Ming-tzu (1935), Akabori (1934), Hrdlička (1932) and Trotter (1934). There are data for four Chinese series but the numbers of bones are small. For each sex considered separately no significant differences are found between the percentage frequencies. For all four series combined the female frequency exceeds the male but the difference is still insignificant. A lack of evidence of distinction is also found in the comparison of the pooled Chinese with the other three Oriental series in the table in the case of male bones. For female bones the percentages show one significant difference: the Japanese value exceeds the Chinese by an amount which is 5·6 times its probable error. These comparisons confirm the conclusion that the frequencies with which septal apertures are found tend to be practically constant for racial populations belonging to the same family of races. The data in the lowest section of Table 4 show that they make some marked distinctions between different families of races. The Oriental, American Negro and Eskimo groups show no significant differences for males and only one for females, but otherwise all the percentages differ with marked significance. The frequency of occurrence of the anomaly is aligned with such a character as skin colour which shows little variation within continental populations but some marked distinctions between such groups.

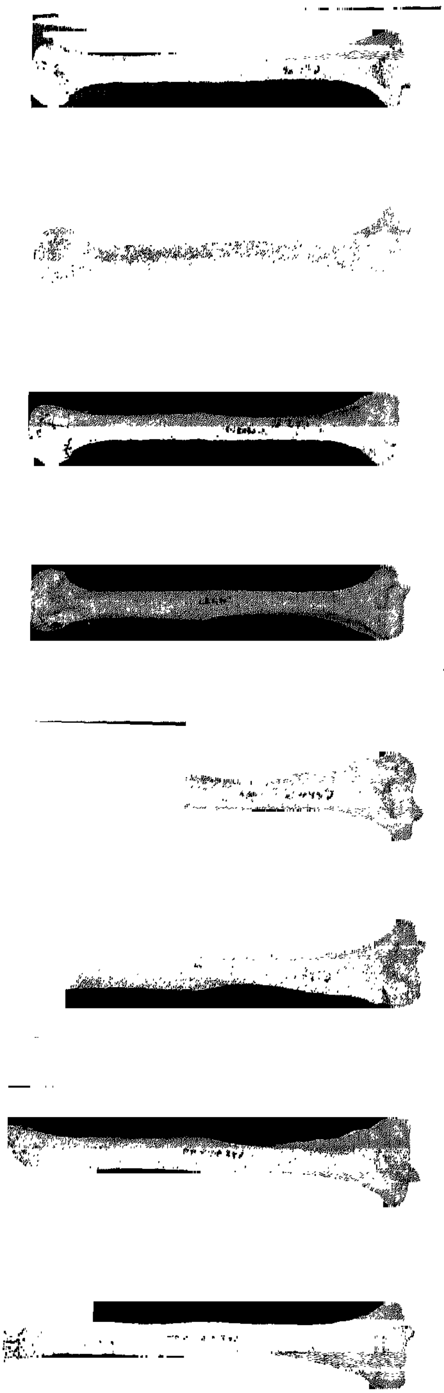
No example of the supracondyloid process was found in either of the Chinese series described in this paper. Black (1925) noted its absence from the prehistoric and modern Chinese humeri which he examined. As far as can be told from the available records, the anomaly is found more frequently in European than in any other populations (Terry, 1930).

* The sizes of the apertures show considerable variation from "pin-points" to holes having a maximum breadth of several mm. Photographs of some of the bones showing the largest apertures are reproduced in Plate I. Only one example of a double aperture was found, this being a female right humerus of the Hsiao T'un series.

EXPLANATION OF PLATE 1.

A. No. 4, ♂, *R* and *L*. Anterior view of typical bones.
B. No. 4, ♂, *R* and *L*. Posterior view of typical bones.
C. No. 26, ♀, *R* and *L*. Anterior view of typical bones.
D. No. 26, ♀, *R* and *L*. Posterior view of typical bones.
E. No. 1, ♂, *R* and *L*, anterior view. Septal aperture in left bone.
F. No. 1, ♂, *R* and *L*, posterior view. Septal aperture in left bone.
G. No. 16, ♀, *R* and *L*, anterior view. Septal aperture in both bones.
H. No. 16, ♀, *R* and *L*, posterior view. Septal aperture in both bones.

ese humeri from Hsiao T'un (all c. one-fifth natural size)



D. C. B. A.

H. G. F. E.

REFERENCES

Akabori, E. (1934). Septal apertures in the humerus in Japanese, Ainu and Koreans. *Amer. J. Phys. Anthrop.* **18**, 395–401.

Black, Davidson (1925). The human skeletal remains from Sha Kuo T'un cave deposit in comparison with those from Yang Shao Tsun and with recent North China skeletal material. *Palaeont. sinica*, series D, **1**, fasc. 3, 1–120.

Haberer, K. A. (1892). *Schädel und Skeletteile aus Peking*. Jena.

Hrdlička, A. (1932). The humerus: septal apertures. *Anthropologie, Prague*, **10**, 31–96.

Kurz, E. (1922). Untersuchungen über das Extremitätenskelett des Chinesen. *Z. ges. Anat.* I. *Z. Anat. EntwGesch.* **66**, 465–557.

Martin, R. (1928). *Lehrbuch der Anthropologie*, 2nd ed. Jena.

Ming-tzu, P'an (1935). Septal apertures in the humerus in the Chinese. *Amer. J. Phys. Anthrop.* **20**, 165–70.

Pearson, K. & Woo, T. L. (1935). Further investigation of the morphometric characters of the individual bones of the human skull. *Biometrika*, **27**, 424–65.

Schreiner, K. E. (1931, 1935). Zur Osteologie der Lappen. *Instituttet for Sammenlignende Kulturforskning*, serie B, *Skrifter*, **18**, 1, 2.

Stevenson, P. H. (1929). On racial differences in stature long bone regression formulae, with special reference to stature reconstruction formulae for the Chinese. *Biometrika*, **21**, 303–21.

Terry, R. J. (1930). On the racial distribution of the supracondyloid variation. *Amer. J. Phys. Anthrop.* **14**, 459–62.

Trotter, M. (1934). Septal apertures in the humerus of American whites and Negroes. *Amer. J. Phys. Anthrop.* **19**, 213–27.

Wagner, K. (1927). Mittelalter-Knochen aus Oslo. *Skr. norske VidenskAkad.* I. Mat.-naturw. Kl. 1926, **7**.

Woo, T. L. (1931). On the asymmetry of the human skull. *Biometrika*, **22**, 324–52.

Woo, T. L. & Morant, G. M. (1932). A preliminary classification of Asiatic races based on cranial measurements. *Biometrika*, **24**, 108–34.

VARIATIONS IN THE WEIGHTS OF HATCHED AND UNHATCHED DUCKS' EGGS

By J. M. RENDEL, *Department of Biometry, University College, London at Rothamsted Experimental Station, Harpenden*

Landauer (1941) has reviewed a great deal of evidence showing that hatchability of the domestic hen's egg depends in part upon its weight. Very large eggs hatch less frequently than medium-sized eggs. The effect of small weight on hatchability is more obscure; it seems probable that small eggs hatch less frequently than medium-sized eggs, but more frequently than large ones. As eggs of extreme weight are less likely to hatch than eggs of medium weight, we may expect to find that the eggs which hatch are less variable in weight than those which fail to do so, and that selection is acting against variation in egg weight. This paper gives an analysis of the variation in weight of the eggs of domestic ducks, and the effect of egg weight on hatchability.

During an experiment on some factors which might influence the weight of table ducklings (Rendel, 1941) some data were collected on the relationship between the weights of eggs which hatched, and the weights of eggs which failed to do so. Two breeds of duck were used in this experiment, a strain of Aylesbury duck which I have labelled N.P.I., and a strain labelled Allports. The two breeds have been treated separately except in the graph, where the figures from the two breeds have been added together.

The weights of the eggs laid by the two groups are given in Table 1. In the N.P.I. group birds labelled 71–82 are in their first laying season, and nos. 248–283 are in their second. In the Allport group nos. 408–416 are in their first season and nos. 208–332 are in their second.

All the eggs recorded here were laid between 26 February and 7 May 1939, a period of 10 weeks. Eggs were collected at the same time every day, and washed immediately. They were allowed an hour to dry before being weighed. As ducks lay their eggs in the morning fairly regularly, insufficient time will elapse between egg laying and weighing to allow an appreciable loss of weight by evaporation. After 1 week's incubation the eggs were examined by candling. All eggs which had not started to develop at all were removed and it was assumed that they had not been fertilized. The remainder were classed as fertile eggs and are those recorded here. It is probable that some eggs which had in fact started to develop, but which had died before the growth of any blood vessels, were classified as infertile. It is unlikely that this error was large or that it would be distributed so as to affect the results in any way. Abnormal eggs, such as double-yolked eggs, shell-less eggs, and very small eggs weighing less than 45 g. with gross deficiencies of yolk or white, were not incubated. There were very few of these.

The Allport and N.P.I. ducks differed considerably from each other. The N.P.I. is a large breed, the Allport is much smaller and, age for age, lays more eggs. The total number of eggs recorded here for N.P.I. ducks is larger than the number recorded for the Allport ducks because there was a higher proportion of young ducks in the N.P.I. group.

The cumulants of the weights of the eggs are given in Table 2. Apart from a slight difference in egg weight, the only major difference between Allport and N.P.I. is in the 4th cumulant, κ_4,

Table 1. *Egg weights*

A. *N.P.I. ducks. Weight of fertile eggs which hatched*

Wt. of eggs in g.	No. of dam 71	72	73	74	75	76	77	78	79	80	81	82	248	252	257	260	261	266	276	279	280	282	283	Total
105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
103	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
102	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
101	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
100	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
99	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
96	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
93	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	2
92	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1
91	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	2
90	.	.	1	.	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.	3
89	.	.	.	.	.	.	.	1	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	2
88	.	.	3	.	.	.	.	.	.	.	.	1	.	.	1	.	.	1	.	.	.	.	.	6
87	.	.	4	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	.	6
86	.	.	.	.	.	.	.	.	1	.	.	.	.	.	2	.	.	.	.	.	.	1	.	4
85	2	.	2	.	.	.	1	.	.	.	1	.	.	.	.	1	.	.	.	1	.	.	.	8
84	2	.	1	.	.	.	.	1	.	.	1	1	.	.	1	.	.	.	.	.	.	1	.	8
83	.	.	2	1	1	1	.	2	.	.	3	1	.	1	1	.	1	.	.	1	.	.	.	15
82	.	.	2	.	1	.	.	.	.	.	1	1	.	1	.	.	.	1	1	1	.	1	.	10
81	1	.	2	.	1	.	1	.	3	.	4	.	.	3	1	.	2	1	.	.	.	.	.	19
80	.	.	.	.	3	.	5	.	1	.	2	1	1	.	.	.	1	2	1	2	.	.	2	21
79	1	.	3	1	2	1	1	2	.	.	.	2	.	1	.	.	.	5	2	2	.	1	3	27
78	3	.	5	.	1	3	7	1	1	.	2	1	1	2	1	2	2	.	1	1	1	.	1	36
77	3	.	1	1	2	2	5	.	.	.	1	.	2	1	.	1	4	1	1	2	.	1	1	29
76	3	.	4	1	3	2	6	.	1	.	1	1	.	1	.	1	4	1	1	1	4	.	3	38
75	2	.	1	.	5	3	6	2	2	.	1	1	1	.	.	4	.	.	3	1	3	.	2	37
74	8	1	4	.	2	2	4	.	2	.	.	.	.	1	.	1	1	1	1	2	4	2	3	39
73	6	2	.	2	1	1	1	1	2	.	.	.	.	.	.	2	1	2	2	2	.	1	2	28
72	5	3	2	.	1	.	3	1	1	1	1	.	2	.	.	.	2	2	2	2	1	.	3	32
71	3	.	3	.	.	1	1	2	4	6	1	.	1	.	.	3	2	.	2	3	.	3	.	35
70	.	4	2	.	1	2	.	1	4	4	.	.	1	.	.	3	1	6	1	3	1	6	3	43
69	.	3	.	.	1	.	.	.	1	6	.	.	3	1	.	2	1	2	3	2	.	2	.	27
68	1	1	1	1	2	2	.	.	2	9	.	.	3	.	.	2	.	1	2	.	2	4	1	34
67	.	3	.	.	.	1	.	1	.	5	.	.	5	.	.	.	1	5	1	.	.	1	.	23
66	.	4	.	.	.	1	.	.	.	5	.	.	2	.	.	2	1	3	1	1	.	.	1	21
65	.	2	.	.	.	6	.	.	3	9	.	.	1	.	.	.	.	.	1	.	3	1	1	27
64	.	2	.	.	.	4	.	.	.	1	.	.	2	.	.	.	.	.	.	.	3	.	2	14
63	.	.	.	.	.	2	.	.	1	1	.	.	1	.	.	.	.	.	.	.	.	.	1	6
62	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	1	.	.	.	.	.	.	.	3
61	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	1
60	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1
59	.	1	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.	3
58	.	.	.	3	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	.	.	5
57	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
55	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
54	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
53	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Total	40	26	43	11	27	35	42	15	30	48	22	10	30	13	13	25	24	34	27	27	23	25	29	619

Table 1 (*cont.*)

B. *N.P.I. ducks. Weight of fertile eggs which did not hatch*

Wt. of eggs in g.	No. of dam																							Total
	71	72	73	74	75	76	77	78	79	80	81	82	248	252	257	260	261	266	276	279	280	282	283	
105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
103	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
102	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
101	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
100	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
99	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
96	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	3
95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5	.	.	.	.	.	.	.	.	5
93	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
92	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	2
91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	2
90	.	.	.	.	.	.	.	.	.	.	1	1	.	1	2	.	.	.	1	.	.	.	.	6
89	.	1	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	2
88	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	1
87	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	1	.	.	.	.	.	.	4
86	1	1	.	.	.	.	.	.	.	.	.	1	.	.	.	1	.	.	1	1	.	.	.	6
85	.	.	.	.	.	.	.	.	.	.	1	.	.	1	2	.	.	.	.	.	.	.	.	4
84	1	.	.	.	1	.	.	.	.	.	2	.	.	1	.	1	.	.	.	1	.	.	.	7
83	1	.	.	2	1	.	1	.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	.	7
82	1	.	1	.	1	.	.	1	.	.	3	2	.	1	.	.	.	.	2	.	.	.	.	12
81	.	1	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	1	1	5
80	2	.	.	.	1	1	5	2	1	.	.	1	.	.	.	.	1	.	.	1	.	.	.	15
79	.	.	.	.	.	1	3	.	.	.	.	.	.	3	2	.	1	1	1	1	2	2	.	17
78	.	.	.	1	1	.	1	.	.	.	.	.	.	.	1	1	.	.	1	1	1	2	.	10
77	2	.	.	.	1	1	2	.	.	.	.	.	.	2	2	1	.	.	.	2	2	1	.	16
76	.	.	.	.	.	.	1	2	2	.	1	1	.	1	.	1	.	1	.	1	2	.	.	13
75	1	.	1	.	.	.	2	.	.	.	1	.	.	1	.	1	.	.	1	.	2	4	.	14
74	.	.	.	1	1	.	1	2	1	.	.	.	1	.	.	.	3	.	.	.	1	1	.	12
73	.	.	1	1	.	.	1	2	.	.	.	.	.	.	.	2	1	.	2	1	4	1	3	19
72	1	.	1	.	.	1	.	.	1	1	1	.	.	.	.	2	1	.	1	2	1	.	2	15
71	3	1	.	1	.	.	.	.	1	.	.	.	.	.	.	4	.	.	.	4	.	3	1	18
70	.	.	1	.	.	.	.	1	1	.	.	.	1	.	.	.	.	.	2	2	2	2	1	13
69	.	1	.	1	.	3	.	1	1	2	1	.	.	.	.	.	1	1	2	.	1	4	1	20
68	.	3	.	.	.	.	.	.	1	.	.	.	1	.	.	.	4	4	.	1	2	1	.	17
67	.	4	.	1	.	2	.	.	.	5	.	.	.	.	.	.	1	3	1	.	.	1	.	18
66	.	3	.	.	1	.	.	.	1	.	.	.	2	.	.	1	.	1	1	.	5	.	.	15
65	.	2	.	1	.	1	.	.	.	1	.	.	.	.	.	.	.	.	.	1	2	.	.	8
64	.	1	.	.	.	.	.	.	1	.	.	.	.	.	.	1	.	.	.	.	2	.	1	6
63	.	2	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1	.	.	3	.	.	7
62	.	1	.	1	.	.	.	.	.	.	.	.	1	.	.	.	3	.	.	.	2	.	.	8
61	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	2
60	.	.	.	1	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	2
59	.	.	.	1	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
58	.	.	.	1	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
57	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
55	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
54	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
53	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
52	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
Total	13	21	5	17	8	12	17	11	11	10	12	13	8	12	22	17	17	12	16	19	35	23	10	341

Table 1 (*cont.*)

C. *Allport ducks. Weight of fertile eggs which hatched*

Wt. of eggs in g.	No. of dam																								Total
	408	409	410	411	412	413	414	415	416	208	209	210	211	213	215	216	219	221	235	238	258	329	331	332	
97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1
96	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1
94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
93	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	1
92	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1
90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.	2
89	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	.	3
88	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	1	3
87	.	.	.	.	.	.	2	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	3
86	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	2	.	.	.	.	.	3
85	.	.	.	.	.	.	1	1	.	.	.	1	.	.	.	.	.	.	2	1	.	.	1	.	7
84	.	.	.	.	.	.	1	.	.	.	.	1	.	1	.	.	.	.	1	1	.	.	1	.	6
83	.	.	.	.	.	.	3	2	.	.	.	.	.	.	.	.	.	.	1	1	.	1	2	.	10
82	.	1	.	.	.	.	2	3	1	1	.	1	1	.	.	.	.	.	.	1	.	.	2	.	13
81	.	5	.	.	.	.	6	1	2	1	.	1	2	.	.	.	.	.	.	1	.	.	1	.	20
80	.	3	.	.	.	1	4	2	.	.	.	2	1	.	.	.	.	.	1	2	.	.	2	.	18
79	.	1	.	.	1	1	3	1	2	1	2	.	.	1	.	1	1	.	.	.	1	.	2	.	18
78	.	5	1	.	1	1	3	2	4	.	1	.	3	.	.	.	1	.	.	.	.	2	3	.	27
77	.	2	1	.	1	1	2	2	4	1	.	.	2	1	1	.	2	.	.	.	.	1	1	1	23
76	.	1	.	.	1	1	5	1	2	4	3	.	4	.	.	1	2	.	.	1	.	.	.	1	27
75	2	3	.	1	3	2	.	5	.	3	3	.	4	2	.	1	2	1	.	.	2	1	1	.	36
74	.	2	2	.	1	1	6	3	1	2	.	.	.	.	1	.	2	1	.	1	.	4	2	2	31
73	1	2	3	1	3	3	1	2	1	1	1	.	2	1	.	2	3	.	.	1	2	2	.	2	34
72	1	1	.	2	.	4	1	1	2	5	1	.	1	2	.	3	1	2	.	4	3	1	.	6	41
71	2	.	.	.	.	3	.	2	3	5	1	.	2	.	.	2	2	1	.	2	4	2	.	3	34
70	.	.	1	4	4	2	1	2	1	2	1	.	2	1	.	6	1	2	.	2	4	.	1	.	37
69	3	.	1	.	1	1	2	.	.	1	.	.	3	1	1	4	.	4	.	.	1	.	.	.	23
68	1	.	.	3	2	4	.	.	1	1	.	.	.	2	1	6	.	6	.	.	2	1	.	3	33
67	4	.	.	.	2	.	.	.	.	.	2	.	.	.	.	7	.	2	.	.	2	4	.	4	27
66	5	.	1	1	1	.	.	.	.	.	2	.	.	2	3	1	.	3	.	1	1	.	.	1	22
65	2	.	.	.	3	6	.	.	.	.	2	.	1	.	3	2	.	1	.	.	.	1	1	2	24
64	2	.	1	.	4	1	.	.	.	.	.	.	.	.	5	4	.	1	.	.	.	1	.	.	19
63	2	.	2	1	1	.	.	.	.	.	2	.	.	.	3	.	.	1	.	.	.	6	.	.	18
62	1	.	1	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4	.	.	8
61	.	.	2	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	4
60	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	3
59	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	1
57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Total	26	26	17	14	31	32	43	30	24	28	22	8	29	14	19	40	17	25	14	20	23	34	21	26	583

Table 1 (*cont.*)

D. *Allport ducks. Weight of fertile eggs which did not hatch*

Wt. of eggs in g.	No. of dam																								Total
	408	409	410	411	412	413	414	415	416	208	209	210	211	213	215	216	219	221	235	238	258	329	331	332	
97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
96	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	2
94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	2
93	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
92	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
91	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4	.	.	.	.	.	4
90	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	2
89	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
88	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1
87	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	3
86	.	.	.	.	.	.	1	1	.	.	.	1	.	.	.	.	.	.	1	.	.	.	.	.	4
85	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	1	.	.	.	1	.	5
84	.	2	.	.	.	.	2	1	.	.	.	.	.	3	.	.	.	.	.	1	.	.	.	.	9
83	.	1	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	1	.	6
82	.	2	.	.	.	3	2	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	2	.	10
81	.	2	.	.	.	2	2	.	.	.	.	.	2	1	.	.	.	.	.	1	.	.	1	.	11
80	.	2	.	.	.	1	1	1	.	.	.	.	.	.	.	.	.	.	1	.	.	.	4	.	10
79	.	3	.	1	.	1	2	.	1	1	1	1	.	1	.	.	.	.	.	1	.	1	1	.	15
78	.	4	.	1	.	.	.	.	3	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	10
77	.	4	2	.	1	1	2	.	2	1	.	.	.	.	.	.	.	.	.	.	1	1	.	.	15
76	.	.	.	2	2	4	.	1	1	1	1	.	1	3	.	.	.	.	.	1	.	3	1	.	21
75	.	4	1	2	1	.	.	2	2	.	.	.	2	1	.	.	.	.	.	3	.	1	.	2	21
74	.	1	.	.	.	2	1	2	1	3	.	.	1	1	.	.	.	.	.	1	.	1	2	1	17
73	1	.	.	1	2	1	1	1	.	1	.	.	.	2	.	2	2	.	.	.	1	1	.	1	17
72	.	.	2	2	1	.	.	1	.	2	2	.	.	2	.	.	.	.	.	1	.	2	.	1	16
71	6	.	3	3	4	1	.	1	.	1	2	.	.	.	.	1	1	.	.	2	2	.	.	.	27
70	4	.	.	.	1	2	.	2	1	.	1	.	.	.	1	.	1	.	.	.	.	3	1	2	19
69	5	.	.	1	1	.	.	.	.	.	.	.	1	.	.	1	.	1	.	.	1	1	1	1	14
68	3	.	1	1	2	1	.	.	.	1	.	.	.	1	.	5	.	1	.	1	1	.	.	2	20
67	1	1	.	3	1	2	.	.	.	1	.	.	1	.	.	2	.	2	.	.	.	.	.	.	14
66	1	.	2	.	1	2	.	.	1	1	.	.	.	.	1	.	.	1	.	.	1	.	.	.	11
65	1	.	.	2	.	.	.	.	.	.	.	.	.	.	.	1	.	2	.	.	.	1	.	1	8
64	2	.	2	1	2	1	.	.	.	.	.	.	.	.	.	2	.	.	.	.	1	1	.	.	12
63	1	.	.	2	.	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	6
62	1	.	1	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	4
61	2	.	1	.	.	.	.	.	.	.	1	.	.	.	.	1	.	.	.	.	.	1	.	.	6
60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	3
59	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	1
57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Total	28	26	16	22	19	25	17	14	12	13	9	3	9	18	3	17	4	8	15	12	8	22	16	11	347

As measured by g_2 or κ_4/κ_3, the value of this cumulant is highly significant for the hatched N.P.I. eggs, but not for the unhatched. In the Allport group neither is significant. There is a considerable difference between hatched and unhatched eggs in both groups. There is a

Table 2

	N.P.I.			Allport		
	Hatched	Unhatched	Diff.	Hatched	Unhatched	Diff.
No. of eggs	619	341		583	347	
Mean weight in g.: κ_1	73·78 ± 0·27	74·17 ± 0·45	0·39 ± 0·35	72·87 ± 0·27	73·66 ± 0·38	0·79 ± 0·32
κ_2	43·87 ± 2·50	69·06 ± 5·30	25·19 ± 3·64	41·32 ± 1·71	49·76 ± 2·67	8·44 ± 2·14
κ_3	133·10	247·70	—	113·50	145·41	—
κ_4	1570·96	533·95	—	172·21	153·76	—
g_1	0·4581 ± 0·0980	0·4316 ± 0·1217	—	0·4273 ± 0·1010	0·4143 ± 0·1308	—
g_2	0·8164 ± 0·1960	0·1120 ± 0·2427	—	0·1001 ± 0·2025	0·0621 ± 0·2612	—

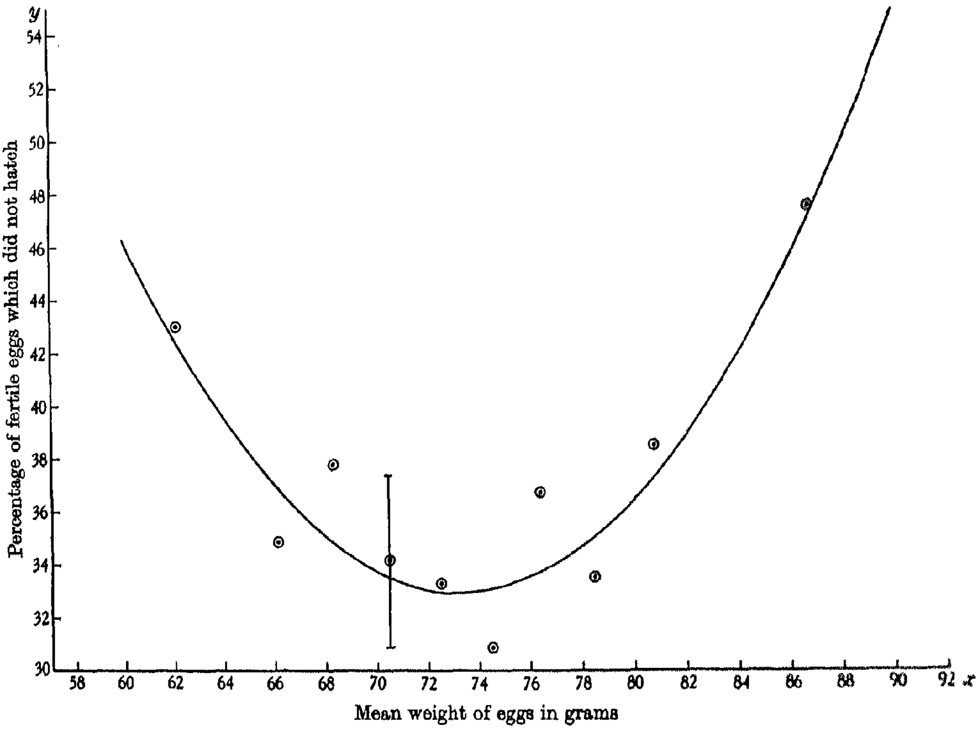


Fig. 1. Combined data for N.P.I. and Allport ducks. The curve is the parabola

$$y = 441\cdot4 - 11\cdot147\,x + 0\cdot07604\,x^2.$$

tendency for the unhatched eggs to be heavier than the hatched and the variance of the former is very much the greater. The distributions of both hatched and unhatched are about equally skew, but the 4th cumulant, κ_4, is greater in the hatched than the unhatched.

These figures show clearly that the further the weight of an egg is from the mean weight, the smaller is its chance of hatching. Egg weight has been plotted against the percentage of eggs which did not hatch in Fig. 1, and a curve fitted. The standard error of the point with the smallest standard error is shown on the graph as a vertical line. The curve shows to what extent egg weight influences hatchability. Eggs with weights outside the range of 70–76 g. lose rapidly in hatchability with each gram, that their weight moves further from the mean. At 67 and 80 g. the loss is approximately 1 % per g., and at 62 and 85 g. the loss is $1\frac{2}{3}$ % per g., with a total loss of 11 % over eggs of mean weight, whereas at 90 g. the rate of loss is $2\frac{1}{2}$ % per g., with a total loss of 21 % over eggs of mean weight.

In breeds of domestic ducks there has been strong selection in favour of a high egg weight. The results given here suggest that it would be advantageous for birds in which a large number of offspring is the first consideration to lay eggs which do not differ widely from the optimum weight for hatchability. It would be interesting to know whether breeds which have not been selected for egg weight by breeders had a variation in egg weight much less than the variation shown here. One would expect, for example, that wild breeds would give a much smaller variance.

Although there are no reliable figures on egg weights, since the time from laying is unknown, there are data in the literature on measurements of length and breadth of eggs of wild birds. From these data it is possible to estimate the amount of variation in egg volume (Appendix 1). The coefficient of variation of various species is given below:

	Coeff. of var. of egg vol.	
English sparrow (*Passer domesticus*)	10·12	Pearson (1901)
Common tern (*Sterna hirundo*)	8·69	Rowan *et al.* (1919)
Common tern (*Sterna hirundo*)	7·75	Watson *et al.* (1923)
Mallard (*Anas platyrhyncha*)	7·18	Fisher (1935)
Common wren (*Troglodytes t.*)	8·13	Fisher (1935)
Greylag goose (*Anser anser*)	10·60	Fisher (1935)
Green woodpecker (*Picus viridis*)	12·9	Fisher (1935)
White-tailed eagle (*Haliaetus albicilla*)	14·2	Fisher (1935)
Cuckoo (*Cuculus canorus*)	14·4	Fisher (1935)

The coefficient of variation of the weights of all eggs considered in this experiment is 9·47. This is well within the limits of variation shown by wild birds' eggs as measured by an estimate of their volume. It is higher than the value found for the wild ducks, this value being 7·18.

There are many factors which influence the egg weight of the duck. Though ducks do lay long uninterrupted series of eggs, they usually lay in clutches varying from two or three eggs to twelve or fifteen with a pause between each clutch. Egg weight is correlated with the position of the egg in the clutch. The first egg is heavier than the last and the intermediate eggs have intermediate weights. Nevertheless, hatchability is not correlated with the position of the egg in the clutch (Rendel, 1941). The time of year is another factor involved. Egg weight rises as the season advances. Then there are certainly some genetic factors. It is of interest therefore to find out whether variation in egg weight due to the differences between individual ducks or variation due to factors which influence the fluctuations of egg weight round the mean of each individual duck, differ significantly for eggs which hatch and eggs which fail to hatch. The former type of variation will be largely genetic, the latter will include many environmental factors, though it is possible that the range of the weights

of eggs laid by a duck is partly determined genetically. An analysis of variance of all good and bad eggs is shown below. The notation is that of Appendix 2.

	Degrees of freedom	Sum of squares	Mean square	
Good eggs: Within progenies	1155	27121	23·481	$\tau^2 = 23{\cdot}481$, variance 0·9547
Between progenies	46	24037	522·54	$\sigma^2 = 19{\cdot}570$, variance 3·9762
Total	1201	51158	42·596	

Bad eggs: Within progenies	641	18459	28·797	$\tau^2 = 28{\cdot}797$, variance 2·5874
Between progenies	46	22237	483·41	$\sigma^2 = 31{\cdot}193$, variance 12·2207
Total	687	40696	59·237	

τ^2 measures the variation which is considered to be largely environmental and σ^2, calculated from the mean square between progenies, measures the variation due to differences between individual ducks, which are considered to be of genetic origin (Appendix 2).

The difference between τ^2 for good and bad eggs is $5{\cdot}316 \pm 1{\cdot}882$ and the difference between σ^2 is $11{\cdot}626 \pm 4{\cdot}025$. Both differences are nearly three times their standard error. We may say therefore that there is selection against ducks which habitually lay heavy or light eggs. This means that there will be selection against certain genotypes and that selection will operate so as to reduce the variation between the mean egg weights of individual ducks in subsequent generations. There is also selection against ducks which lay eggs with a very wide range of egg weights. It does not necessarily follow that such selection will have any effect on future generations, though it may do so.

In the light of this result we may review some other similar results. Crampton (1904) found that the pupae of *Philosamia cynthia* which died were significantly more variable than the survivors as regards seven measurements. Weldon (1901) in *Clausilia laminata* and Di Cesnola (1907) in *Helix arbustorum* found that the shells of young individuals varied more than the shells of old ones. In such cases it was not clear how far the variation was due to nature or nurture. It is possible that the individuals differing most from the mean had been exposed to more extreme conditions and that these conditions brought about the greater variability as well as the higher mortality. It has been shown, for example, that the weights of heart, liver and kidneys when diseased are more variable than the weights of these organs when healthy, and the correlation between the weights of these organs is less when they are diseased (Greenwood, 1904). This increase in variability and decrease in correlation is thought to be due to reaction of the body to disease. If this is the case, death of the more variable individuals has no selective influence. In the case of the ducks' eggs it has been possible to eliminate the day-to-day variation in egg weight and therefore a good deal of the variation due to nurture. We can at least say that there is some selection against ducks which habitually lay very large or very small eggs.

Summary

A description of the variation of the weights of eggs laid by forty-seven ducks is given. Eggs which hatch are compared to eggs which do not hatch. Variation in egg volume of some wild species of birds is compared with variations in egg weight of the domestic duck. It is concluded that selection will tend to reduce variation in egg weight.

I should like to thank Dr Helen Spurway for preparing the graph, and Prof. J. B. S. Haldane for his help in the mathematical treatment of the results and for contributing the two appendices.

References

Cesnola, A. P. Di (1907). *Biometrika*, **5**, 387–99.
Crampton, H. E. (1904). *Biometrika*, **3**, 113–30.
Fisher, R. A. (1935). *Proc. Roy. Soc.* B, **122**, 1–26.
Greenwood, M. (1904). *Biometrika*, **3**, 63–83.
Landauer, W. (1941). *Bull. Storrs Agric. Exp. Sta.* no. 236.
Pearson, Karl (1901). *Biometrika*, **1**, 256–7.
Rendel, J. M. (1941). *Emp. J. Exp. Agric.* **9**, 50–7.
Rowan, W. *et al.* (1919). *Biometrika*, **12**, 308–54.
Watson, D. M. S. *et al.* (1923). *Biometrika*, **15**, 294–345.
Weldon, W. F. R. (1901). *Biometrika*, **1**, 109–24.

Appendix 1. THE COEFFICIENT OF VARIATION OF EGG VOLUME

By J. B. S. HALDANE, F.R.S.

A number of workers have measured the length L and breadth B of birds' eggs and found them to be more or less normally distributed. All authors give the means $\bar{L}$ and $\bar{B}$. Pearson and his colleagues give the mean value $\bar{I}$ of the index $I = B/L$, which is also nearly normally distributed, and the coefficients of variation l, b and i, of L, B and I. Fisher gives the variances λ and β of the length and breadth, and their covariance κ.

It is required to find the coefficient of variation v of the volume V, supposing that $V = kLB^2$, where k is a constant. That is to say the eggs are all supposed to be of the same shape apart from changes of scale in L and B. We also suppose L and B to be normally correlated with coefficient ρ. Let $u^2 = 2\rho bl$.

Then $L = \bar{L}(1+lx)$, and $B = \bar{B}(1+by)$, where x and y are correlated reduced normal variates (with zero mean and unit variance). The means of odd powers and products of x and y vanish, and

$$\overline{x^2} = \overline{y^2} = 1, \quad \overline{xy} = \rho, \quad \overline{x^4} = \overline{y^4} = 3, \quad \overline{x^3y} = \overline{xy^3} = 3\rho, \quad \overline{x^2y^2} = 1+2\rho^2, \quad \overline{x^4y^2} = 3(1+4\rho^2).$$

$$V = k\bar{L}(\bar{B})^2(1+lx)(1+by)^2.$$

Hence
$$\bar{V} = k\bar{L}(\bar{B})^2(1+2lb\overline{xy}+b^2\overline{y^2}) = k\bar{L}(\bar{B})^2(1+b^2+u^2),$$

$$\overline{V^2} = \bar{V}^2(1+v^2) = k^2(\bar{L})^2(\bar{B})^4[1+l^2+4u^2+6b^2+3(2l^2b^2+u^4+4b^2u^2+b^4)+3b^2(l^2b^2+u^4)].$$

So
$$v^2 = l^2+2u^2+4b^2+2(2l^2b^2-l^2u^2-u^4-5u^2b^2-3b^4)+\dots.$$

Since l, b, and u are of the order of 0·05, we may take

$$v = \sqrt{(l^2+2u^2+4b^2)}, \tag{1}$$

with an error of the order of 4 %, which is generally negligible.

Since $\lambda = (\bar{L})^2 l^2, \quad \beta = (\bar{B})^2 b^2, \quad \kappa = \rho \bar{L}\bar{B} l b = \tfrac{1}{2}\bar{L}\bar{B}u^2,$

$$v = \sqrt{\left(\frac{\lambda}{(\bar{L})^2} + \frac{4\kappa}{\bar{L}\bar{B}} + \frac{4\beta}{(\bar{B})^2}\right)}. \tag{2}$$

If i is given, but not κ, we proceed as follows:

$$I = \frac{\bar{B}}{\bar{L}}(1+by)(1+lx)^{-1}.$$

This may be expanded in a series, and the moments calculated. Theoretically they are all infinite, since if L is normally distributed it can be zero. However, the series for the moments are asymptotic expansions, and the first few terms give good approximations:

$$\bar{I} = \frac{\bar{B}}{\bar{L}}\left[1 + (l - \rho b)\sum_{r=1}^{\infty} \frac{l^{2r-1}(2r!)}{2^r r!}\right] = \frac{\bar{B}}{\bar{L}}[1 + (l^2 - \tfrac{1}{2}u^2)(1 + 3l^2 + 15l^3 + \ldots)],$$

$$\overline{I^2} = \left(\frac{\bar{B}}{\bar{L}}\right)^2 + \bar{I}^2 i^2 = \left(\frac{\bar{B}}{\bar{L}}\right)^2 [1 + 3l^2 - 2u^2 + b^2 + 3(5l^4 - 4l^2u^2 + 3l^2b^2 + 3u^4) + \ldots],$$

$$i^2 = l^2 + b^2 - u^2 + \text{etc.},$$

and
$$v = \sqrt{(3l^2 + 6b^2 - 2i^2)}, \tag{3}$$

again with an error of the order of about 4 %.

Appendix 2. INTERPRETATION OF THE GREATER VARIANCE OF THE UNHATCHED EGGS

By J. B. S. HALDANE, F.R.S.

First let us consider those eggs which hatched. Let the weight of such an egg be $W+w$, where W is the true mean, or the mean of a very large sample. Let $w = x\sigma + y\tau$, where x and y are uncorrelated reduced normal variables, i.e. $\bar{x} = \bar{y} = 0$, $\overline{x^2} = \overline{y^2} = 1$; and let x be constant for any given duck. That is, σx represents the deviation from mean egg weight due to a given duck's make-up, and τy the deviation of the egg weight from $W + \sigma x$ during the duck's life. We wish to estimate σ and τ, and to know whether they differ significantly from σ' and τ', the corresponding quantities, for unhatched eggs.

Let there be n ducks laying N good eggs in all. Let the rth duck lay k_r good eggs, so that $N = \sum_{r=1}^{n} k_r$.

Then the total sum of squares of deviations

$$\Sigma(w - \bar{w})^2 = \left(\frac{N^2 - \Sigma k_r^2}{N}\right)\sigma^2 + (N-1)\tau^2,$$

and the corresponding variance (mean square) is $\dfrac{N^2 - \Sigma k_r^2}{N(N-1)}\sigma^2 + \tau^2$, or $\left[1 + \dfrac{N - \Sigma k_r^2}{N(N-1)}\right]\sigma^2 + \tau^2$, which is nearly $\sigma^2 + \tau^2$. The sum of squares of deviations within progenies of individual ducks $\Sigma(w - \bar{w}_r)^2 = (N-n)\tau^2$, and the corresponding mean square is τ^2. Subtracting the two sums of squares, the sum of squares of deviations between progenies is

$$(\Sigma\bar{w}_r - \bar{w})^2 = \frac{N^2 - \Sigma k^2}{N}\sigma^2 + (n-1)\tau^2,$$

and the mean square is
$$\frac{N^2 - \Sigma k^2}{N(n-1)}\sigma^2 + \tau^2.$$

We have therefore:

	Degrees of freedom	Sum of squares	Mean square
Within progenies	$N-n$	$(N-n)\,\tau^2$	τ^2
Between progenies	$n-1$	$\dfrac{N^2-\Sigma k^2}{N}\,\sigma^2+(n-1)\,\tau^2$	$\dfrac{N^2-\Sigma k^2}{N(n-1)}\,\sigma^2+\tau^2$
Total	$N-1$	$\dfrac{N^2-\Sigma k^2}{N}\,\sigma^2+(N-1)\,\tau^2$	$\dfrac{N^2-\Sigma k^2}{N(N-1)}\,\sigma^2+\tau^2$

The sum of squares of deviations between means of progenies, and the total, are readily found, and the variance τ^2 within progenies is found by subtraction. σ^2 is then readily found.

THE PROBABILITY INTEGRAL FOR TWO VARIABLES

By C. NICHOLSON, M.C., M.A., M.D.

1. Introduction

A geometrical approach to problems connected with the normal bivariate surface (Nicholson, 1941) suggested a method for the direct integration of the surface which could be used to calculate a table which might be simpler in use than the present table of d/N calculated by Everitt (1912), Lee (1915), Lee (1927), and Elderton, Moul, Fieller, Pretorius & Church (1930) and republished in *Tables for Statisticians and Biometricians*, Part II. The results of this inquiry are here presented, and I must acknowledge the assistance which I have received in the preparation of the paper from Prof. E. S. Pearson and Mr N. L. Johnson.

A brief recapitulation of the relevant results from the earlier paper will first be given. Fig. 1 shows diagrammatically an elliptic contour of the normal correlation surface, for

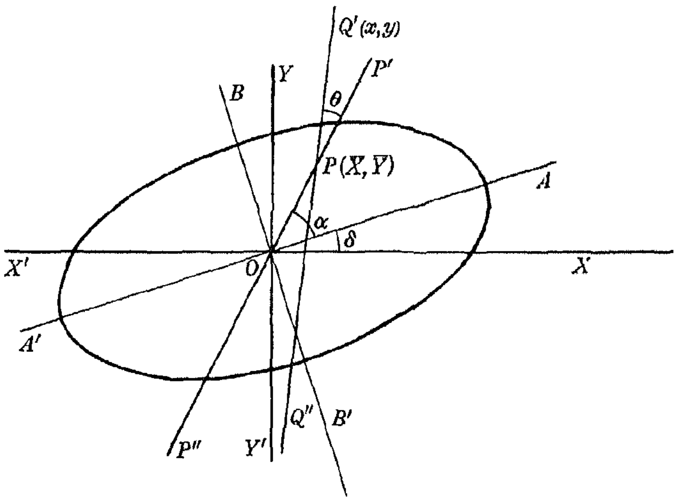


Fig. 1. Projection of normal bivariate surface to illustrate the geometry of the ratio between variables.

which the mean is at O, and the standard deviations and correlation coefficient of the variables are σ_x, σ_y, and r respectively. The principal axes make angles of δ and $\frac{1}{2}\pi+\delta$ with OX, and the standard deviations in these principal directions are a and b. If P is a point $(\bar{X}, \bar{Y})$, then the ratio of OP to the standard deviation of the plane section of the normal surface through OP is

$$p = \frac{1}{\sqrt{(1-r^2)}}\sqrt{\left\{\frac{\bar{X}^2}{\sigma_x^2} - \frac{2r\bar{X}\bar{Y}}{\sigma_x\sigma_y} + \frac{\bar{Y}^2}{\sigma_y^2}\right\}} \tag{1}$$

OP makes an angle α with the major axis of the ellipse. The earlier paper then considered the distribution of the ratio

$$v = \frac{y-\bar{Y}}{x-\bar{X}} = \tan(\alpha+\theta+\delta), \tag{2}$$

which is constant along $Q'PQ''$. The frequency distribution of v may clearly be derived

from that of θ, and it was shown that the distribution of θ could most simply be obtained from that of an angle ϕ given by

$$\phi = \tan^{-1}\{(a/b)\tan(\alpha+\theta)\} - \tan^{-1}\{(a/b)\tan\alpha\}. \tag{3}$$

Geometrically ϕ is the angle between lines corresponding to $P'OP''$ and $Q'PQ''$ after a transformation which substitutes for the correlation surface with its elliptical contours an equivalent system with circular contours having a common standard deviation for every section of $\sqrt{\left(\frac{a^2+b^2}{2}\right)}$, that is, $\sqrt{\left(\frac{\sigma_x^2+\sigma_y^2}{2}\right)}$. As a consequence of this transformation it follows that the cumulative frequency of θ and therefore of $\tan^{-1}v-(\alpha+\delta)$ is that of ϕ, or, as was proved,

$$P\{0 \leqslant \theta \leqslant \Theta\} = \int_0^{\Phi} \frac{e^{-\frac{1}{2}p^2}}{\pi}\left\{1 + p\cos\phi\, e^{\frac{1}{2}(p\cos\phi)^2}\int_0^{p\cos\phi} e^{-\frac{1}{2}x^2}\,dx\right\}d\phi, \tag{4}$$

where p and Φ are given in terms of Θ and the constants of the surface by equations (1) and (3). It was further shown that (4) may be expanded into a form

$$\phi/\pi + 2V(p,\phi) \tag{5}$$

where
$$V(p,\phi) = \frac{\sin\phi}{2\pi}\{A_0\cos\phi + A_1\cos^3\phi + A_2\cos^5\phi + \dots\}, \tag{6}$$

and
$$A_0 = \int_0^p pe^{-\frac{1}{2}p^2}\,dp, \tag{7}$$

$$A_1 = \frac{1}{1\cdot 3}\int_0^p p^3 e^{-\frac{1}{2}p^2}\,dp, \tag{7a}$$

....

$$A_n = \frac{1}{1.3.5\dots(2n+1)}\int_0^p p^{(2n+1)}\, e^{\frac{1}{2}-p^2}\,dp. \tag{7b}$$

Now if we refer to Fig. 3 and suppose that the lines $P'OP''$, and $Q'PQ''$ of Fig. 1 have been transformed into $P'OP''$, and $H'PH''$, the cumulative frequency is the content of the double sector between the planes $H'PH''$ and $P'PP''$, and this is equal to the content of the double sector between the planes $Y_hOY'_h$ and $P'OP''$, that is to ϕ/π, together with twice the content above the triangle OPH, so that $V(p,\phi)$ is the content of this triangle.

2. The standardized surface

We may now apply these conclusions to the special case of the standardized normal surface

$$z = \frac{N}{2\pi\sqrt{(1-r^2)}}\exp\left[-\frac{1}{2}\frac{x^2-2rxy+y^2}{1-r^2}\right],$$

which, using the nomenclature of *Tables for Statisticians and Biometricians*, is divided into four parts, a, b, c, and d by the two planes $x = h$, and $y = k$ parallel to the co-ordinate axes and intersecting in the ordinate at h, k. If we further subdivide this surface by a plane $P'OP''$ $(x/h = y/k)$ through the ordinates at the origin and at h, k we have the position in Fig. 2, and it will be appreciated that the content between the planes $P'PP''$ and $H'PH''$, as also the content between the planes $P'PP''$ and $K'PK''$ may be calculated separately by means of (4) (with the assumption that both deviations are positive), and that the sum of these contents is

$$\phi_h/\pi + 2V(p,\phi_h) + \phi_k/\pi + 2V(p,\phi_k) = (a+d)/N. \tag{8}$$

To do this we must, as before, refer the distribution of Fig. 2 to the equivalent symmetrical distribution obtained by transformation and illustrated in Fig. 3. Here

$$p = \sqrt{\left\{\frac{h^2 - 2rhk + k^2}{1 - r^2}\right\}}, \tag{1a}$$

and, since $\sigma_x = \sigma_y$, the angle which the major and minor axes make with the primary co-ordinate axes is $\delta = \frac{1}{4}\pi$, so that

$$\alpha = \tan^{-1}(k/h) - \tfrac{1}{4}\pi = \tan^{-1}\{(k-h)/(k+h)\}. \tag{9}$$

The standard deviations on the major and minor axes of the ellipse are respectively $\sqrt{(1+r)}$ and $\sqrt{(1-r)}$. The angular deviation θ_h ($P'PH'$ in Fig. 2) is $\frac{1}{4}\pi - \alpha$, so that from (3)

$$\phi_h = \tan^{-1}\left\{\frac{\sqrt{(1+r)}}{\sqrt{(1-r)}}\right\} - \tan^{-1}\left\{\frac{(k-h)\sqrt{(1+r)}}{(k+h)\sqrt{(1-r)}}\right\} = \tan^{-1}\left\{\frac{h\sqrt{(1-r^2)}}{k-rh}\right\}. \tag{10}$$

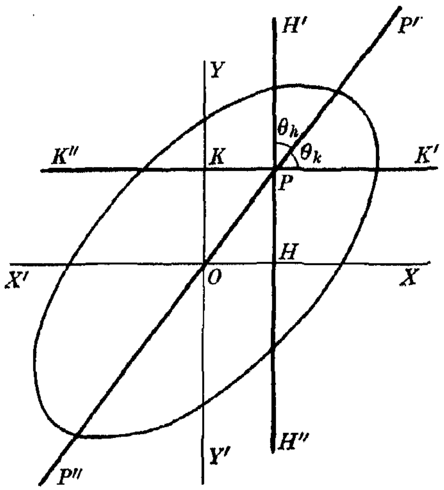


Fig. 2.

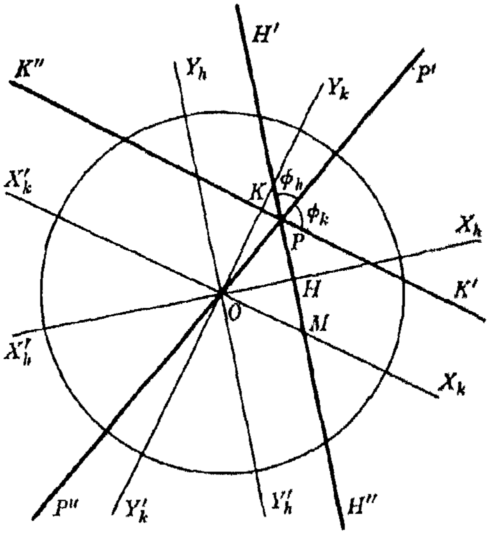


Fig. 3.

Fig. 2. Projection of the standardized normal surface $r = 0{\cdot}60$.

Fig. 3. Projection of the symmetrical surface equivalent to that illustrated in Fig. 2 to show the relationship between θ and ϕ.

Similarly, $\theta_k = \frac{1}{4}\pi + \alpha$, and

$$\phi_k = \tan^{-1}\left\{\frac{\sqrt{(1+r)}}{\sqrt{(1-r)}}\right\} + \tan^{-1}\left\{\frac{(k-h)\sqrt{(1+r)}}{(k+h)\sqrt{(1-r)}}\right\} = \tan^{-1}\left\{\frac{k\sqrt{(1-r^2)}}{h-rk}\right\}. \tag{10a}$$

It will be seen that for any value of r, $\phi_h + \phi_k$ is constant at all points in the distribution depending only on r; in fact

$$\phi_h + \phi_k = \kappa = 2\tan^{-1}\left\{\frac{\sqrt{(1+r)}}{\sqrt{(1-r)}}\right\} = \pi - \cos^{-1} r. \tag{11}$$

Also if we consider the triangle OPH in Fig. 3 we have the following identities:

$$\angle OPH = \phi_h = \tan^{-1}\left\{\frac{h\sqrt{(1-r^2)}}{k-rh}\right\}, \qquad OH = h,$$

$$OP = p = \sqrt{\left\{\frac{h^2 - 2rhk + k^2}{1-r^2}\right\}}, \qquad PH = q_h = PM - HM = \frac{k - rh}{\sqrt{(1-r^2)}}.$$

This is the position reached by Sheppard (1900) by another line of reasoning, and he goes on to discuss the calculation of a table of V and its application to practical work. Almost contemporaneously Pearson (1901) published a method for integrating the surface as a polynomial in r with coefficients which are functions of h and k, the tetrachoric functions of Everitt (1910). This method was devised primarily for the calculation of the correlation coefficient for a fourfold table which was supposedly normal in distribution, but the polynomial converges so slowly for high values of r that it was never very satisfactory. In its place a table of d/N was calculated for high positive values of r (Everitt, 1912) and for high negative values (Lee, 1915). In Fig. 3 since

$$PH = \frac{k-rh}{\sqrt{(1-r^2)}}$$

for all values of h, it will be seen that geometrically

$$d/N = \frac{1}{2\pi}\int_h^\infty e^{-\frac{1}{2}x^2}\int_{\frac{k-rx}{\sqrt{(1-r^2)}}}^\infty e^{-\frac{1}{2}y^2}\,dy\,dx,$$

and from this double integral the tables were calculated by quadrature. The table was later (Lee, 1927; Elderton *et al.* 1930) extended to all values of r positive and negative at intervals of 0·05. This is a very extensive table running to more than 20,000 entries; moreover, it is a table of three arguments demanding a not very satisfactory triple interpolation. It is suggested that the table of V given at the end of this paper, of two arguments and extending to no more than 900 entries, would give at least equal accuracy, and for many purposes would be as convenient to use.

3. Calculation of the table

For such a table of V the arguments must obviously be chosen from the sides and angles of the triangle OPH; the side OP may be at once ruled out as requiring far too much preliminary calculation, and for the same reason the angle ϕ or any of its functions is not suitable; this leaves us with the two sides

$$OH = h, \quad \text{and} \quad PH = q = \frac{k-rh}{\sqrt{(1-r^2)}}.$$

While these are suitable arguments, the formula (5) is not suited to them; if, however, we integrate the content of the triangle from the origin we have, using linear variables,

$$V(h,q) = \frac{1}{2\pi}\int_0^h e^{-\frac{1}{2}x^2}\int_0^{qx/h} e^{-\frac{1}{2}y^2}\,dy\,dx, \tag{12}$$

or, using an angular variable,

$$V(h,q) = \frac{1}{2\pi}\int_0^{(\frac{1}{2}\pi-\phi)} \{1-e^{-\frac{1}{2}(h\sec\psi)^2}\}\,d\psi, \tag{13}$$

and these may be expanded into a form

$$V(h,q) = \frac{1}{2\pi}\left\{B_0\frac{q}{h} - B_1\frac{q^3}{3h^3} + B_2\frac{q^5}{5h^5} - B_3\frac{q^7}{7h^7} + \dots\right\}, \tag{14}$$

where
$$B_0 = \int_0^h h e^{-\frac{1}{2}h^2} dh, \tag{15}$$

$$B_1 = \frac{1}{2}\int_0^h h^3 e^{-\frac{1}{2}h^2} dh, \tag{15a}$$

....

$$B_n = \frac{1}{2^n . n!}\int_0^h h^{(2n+1)} e^{-\frac{1}{2}h^2} dh. \tag{15b}$$

The first requisite for calculating V is, then, a table of B_n and

$$B_n = 1 - e^{-\frac{1}{2}h^2}\left\{1 + \frac{h^2}{2} + \frac{h^4}{2^2 . 2!} + \frac{h^6}{2^3 . 3!} + \dots + \frac{h^{2n}}{2^n . n!}\right\}. \tag{15c}$$

The value actually tabled was $B_n/2\pi(2n+1)$, and since (when $q = h$)

$$\frac{1}{2\pi}\{B_0 - \tfrac{1}{3}B_1 + \tfrac{1}{5}B_2 - \tfrac{1}{7}B_3 + \dots\} = \frac{1}{2}\left\{\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2} dh\right\}^2, \tag{16}$$

there is a very useful check on accuracy. This table was taken to 8 places of decimals and the figures for $h = 3$ are given to show the number of terms which are necessary for the higher values of h.

Table 1

n	$B_n/2\pi(2n+1)$	n	$B_n/2\pi(2n+1)$
0	0·15738689	1	0·04981022
2	2630583	3	1495384
4	827423	5	429819
6	206840	7	91871
8	37689	9	14318
10	5054	11	1664
12	513	13	149
14	41	15	11
16	3	17	1
Sum	0·19446835	Sum	0·07014239
	0·07014239		
	$V(3, 3) = 0·12432596$		
	$2V(3, 3) = 0·24865192$		
	$\sqrt{\{2V(3, 3)\}} = 0·49865010$		

$B_n/2\pi(2n+1)$ was then divided by $h^{(2n+1)}$ and the result multiplied by the successive values of $q^{(2n+1)}$; for this purpose it was not necessary to use all the figures of the powers except in the cases of B_0, B_1, and B_2. Beyond that a diminishing number of figures is enough to maintain accuracy in the 8th decimal place.

It is obvious that this method can only be used when $q < h$, so that a diagonal half of the table of V was calculated in this way; for the other half the identity

$$V(h, q) + V(q, h) = \frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2} dh \times \frac{1}{\sqrt{(2\pi)}}\int_0^q e^{-\frac{1}{2}q^2} dq \tag{17}$$

was used. The 8th place of decimals was now discarded and the remaining 7-figure table checked for accuracy in both directions by fourth or even in some cases fifth differences. When it was proved that the error in the 7th place was not greater than 1, the 7th place was discarded and the table completed.

The slope of the table and the magnitude of differences may be gathered from the two differential equations
$$\frac{\partial V}{\partial q} = \frac{1}{2\pi}\frac{h}{h^2+q^2}\{1-e^{-\frac{1}{2}(h^2+q^2)}\}, \tag{18}$$
and
$$\frac{\partial V}{\partial h} = \frac{1}{2\pi}\left\{e^{-\frac{1}{2}h^2}\int_0^q e^{-\frac{1}{2}q^2}dq - \frac{q}{q^2+h^2}(1-e^{-\frac{1}{2}(q^2+h^2)})\right\}. \tag{19}$$
In the first case, first differences while large do not change much within the range of the table, so that second differences are not appreciable. In the second case, first differences are greater at first and moreover change sign within the range of the table, so that second differences are not negligible except for high values of h.

All use of the table demands the measure of an angle (expressed as a fraction of π), and its trigonometric functions; the auxiliary table (Table 7) was therefore added, and, bearing in mind that the main use of the table must be to obtain a value for the κ of equation (11) from r, the argument chosen was r. From r the succeeding columns of the table were calculated in the order given, the angle being calculated as the inverse tangent; the angle was checked by differences and, as an additional check, was calculated by inverse interpolation from a 7-figure table of the sine. Interpolation is satisfactory in all columns of the table as far as $r = 0{\cdot}80$; intermediate values are not often required above this and can usually be obtained by interpolating for $\frac{1}{2}\pi - \lambda$.

4. Notes on the table

The relationship of $V(h, q)$ to d/N is given by the equation
$$V(h, q_h) + V(k, q_k) + \kappa/2\pi = d/N + \frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}dh + \frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^k e^{-\frac{1}{2}k^2}dk. \tag{20}$$
Although $V(h, q)$ has been described as the content above the triangle OPH this is not strictly true; when the point P in Fig. 3 lies between the lines $X_hOX'_h$ and $X_kOX'_k$, rh is greater than k so that q, and so also V, is negative, so that strictly speaking V is not a measure of volume but a mathematical conception. As with d/N in the quadrants where h and k are of opposite sign, r is taken to be negative with the deviates both positive. When r is negative κ is replaced by its complement $\pi - \kappa$.

The limiting values of V, when h and q are beyond the range of the table, are necessary for the fitting of a surface.

(1) When $h = 0$, then $B_n = 0$ for all n, and
$$V(0, q) = 0.$$
(2) When h is finite and

(a) $k = rh$, then $q = 0$, and
$$V(h, 0) = 0,$$
(b) $k = \infty$, then $q = \infty$, and
$$V(h, \infty) = \frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}dh. \tag{21}$$
(3) When $h = \infty$ and

(a) k is finite, then $B_n = 1$, and $\dfrac{q}{h} = \dfrac{-r}{\sqrt{(1-r^2)}}$ so that
$$V(\infty, q) = \frac{1}{2\pi}\left\{\tan^{-1}\left(\frac{-r}{\sqrt{(1-r^2)}}\right)\right\} = \frac{1}{2\pi}(\tfrac{1}{2}\pi - \kappa), \tag{22}$$
(b) $k = \infty$, it is not possible to assign a value to $V(h, q_h)$ or $V(k, q_k)$ separately, but their sum
$$V(h, q_h) + V(k, q_k) = \frac{1}{2\pi}(\pi - \kappa). \tag{23}$$

It must sometimes happen, especially when the value of r is high, that q lies outside the range of the table; in such cases we may make use of the identity

$$V(h,q) = \frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}dh - \frac{1}{2\pi}\tan^{-1}(h/q) + R, \tag{24}$$

where R, geometrically is the content of the sector $H'PP'$ in Fig. 3, and when q is greater than 3 the value of R is negligible. When $q = 3$, the value of R does not exceed 0·0000288, when $q = 4$ R does not exceed 0·0000005.

5. USES OF THE TABLE

There are three main uses of the table.

A. It may be used to calculate the probability integral for the distribution of the ratio between normal variables, v. From the given distribution values are calculated for p as in (1), for the ratio between the principal axes, a/b, and for $\tan\delta = t$. Then taking $\overline{Y}/\overline{X}$ as v_0,

$$\tan\alpha = \frac{v_0 - t}{1 + v_0 t} \tag{25}$$

and

$$\tan(\alpha + \theta_n) = \frac{v_n - t}{1 + v_n t}, \tag{25a}$$

so that it is easy to get out a series of values of ϕ_n by equation (3), corresponding to a series of values of v_n, with the aid of Table 7. The table of V is then entered with $h = p\sin\phi$, and $q = p\cos\phi$, and

$$P\{v_0 \leqslant v \leqslant v_n\} = \phi_n/\pi + 2V(p\sin\phi_n,\ p\cos\phi_n). \tag{26}$$

B. For most uses of the table it is of course necessary to evaluate both $V(h, q_h)$, and $V(k, q_k)$; it is the sum of these which we may designate $W(h, k, r)$ which is to be used. In obtaining a correlation coefficient from a fourfold table of supposedly normal distribution we must use the equation

$$(a+d)/N = \kappa/\pi + 2W(h,k,r). \tag{27}$$

Inverse interpolation is tedious, but

$$\frac{d\{(a+d)/N\}}{dr} = \frac{1}{\pi\sqrt{(1-r^2)}}e^{-\frac{1}{2}p^2} \tag{28}$$

which may readily be calculated as

$$\frac{2}{\sqrt{(1-r^2)}} \times \frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}h^2} \times \frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}q^2}, \tag{28a}$$

so that having obtained an approximate solution less than the true solution it is possible to obtain further accuracy by the use of the Newton-Raphson method of solving equations. It should be noted that most of the value of $(a+d)/N$, when h and k are not large, is carried by κ/π.

C. The main use of the table is no doubt the fitting of a normal surface to a given distribution, and the lay-out for this purpose is illustrated in Fig. 4. It will be seen that $W(h, k, r)$, being the content of the quadrilateral $OHPK$, may be used in much the same way as d/N for obtaining the content of any cell; thus the cell $P_{12}P_{13}P_{17}P_{18}$ is given by $W_{13}+W_{17}-W_{12}-W_{18}$, the cell $P_{13}P_{14}P_{18}P_{19}$ by $W_{18}+W_{19}-W_{13}-W_{14}$, the cell containing the origin is given by the sum of the W's for the surrounding four nodes. It will be seen that in the case of a 4×4-fold table it is necessary to get values of $W(h, k, r)$ for 25 nodes, but of these the table of $V(h, q)$ is used for 9 only, the remainder are functions of $\kappa/2\pi$ and of the

probability integrals for h and k, so that 18 double interpolations are required. In the case of the table of d/N, again 25 values of d/N are to be calculated, but again 9 only of these demand triple interpolation, most of the remainder demand double interpolation. On the whole, then, the table of $V(h, q)$ has no great advantage over the table of d/N except that it is so much more compact.

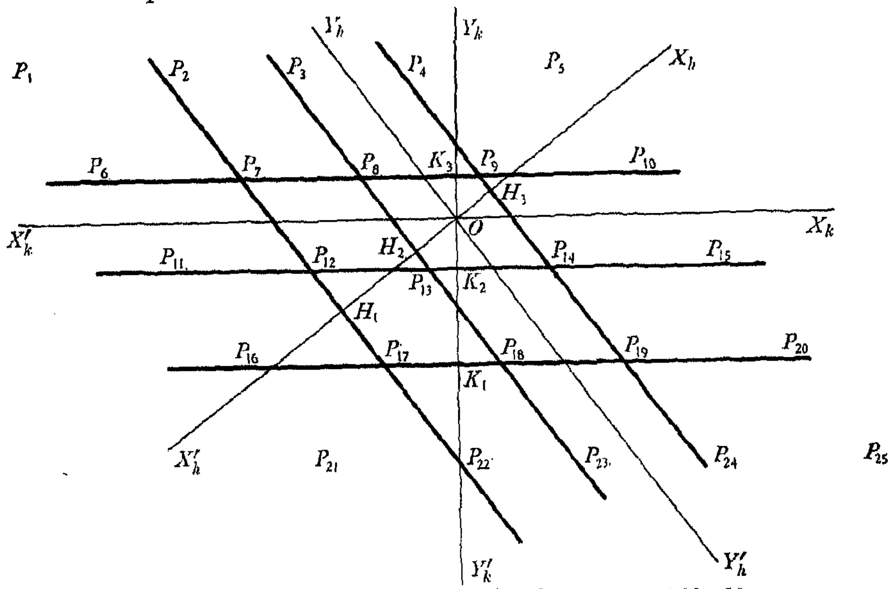


Fig. 4. Lay-out for the fitting of a normal surface to a 4×4-fold table.

6. Example*

The classification of the female pelvis according to the shape of the brim (illustrated in Fig. 5) has hitherto depended on the pelvic index, the percentage ratio of the antero-posterior to the greatest transverse diameter, $100AB : CD$, following Turner (1886). Caldwell & Moloy (1938) have reintroduced a second criterion of classification depending on the relative position on the antero-posterior diameter of the point where it is crossed by the transverse diameter; this may be described by the sagittal index, the percentage ratio of the posterior part to the whole antero-posterior diameter, $100AO : AB$. Measurements in 329 cases made by an accurate technique of X-ray pelvimetry (Nicholson, 1936) give the following constants for these two indices:

	Mean	Standard deviation
Pelvic index	89·92	8·63
Sagittal index	34·69	4·44

As with other anatomical measurements, the distribution of these indices fits well to the normal curve; this may be appreciated from the marginal frequencies in Table 2. The correlation coefficient between them is 0·40. Now it is claimed that a low value of the

* The figures used in this example are taken from an inquiry into the value of X-ray pelvimetry in obstetrics which is assisted by a grant from the Medical Research Council.

sagittal index is due to an imbalance of the sex hormones with male predominance, and Caldwell & Moloy rather beg the question by naming the pelvis of this type the 'android' pelvis. The normality of the distribution of the sagittal index is already a strong argument against this theory, which would be reinforced if it were proved that the correlation between the indices was normal correlation. Further, it is necessary to establish this normality before discussing the frequency of any of Caldwell & Moloy's types. The 329 cases are arranged below in a 4 × 4-fold table (Table 2), the figures in brackets are the theoretical normal frequencies and the procedure used in calculating these is shown in the succeeding Tables, 3–5.

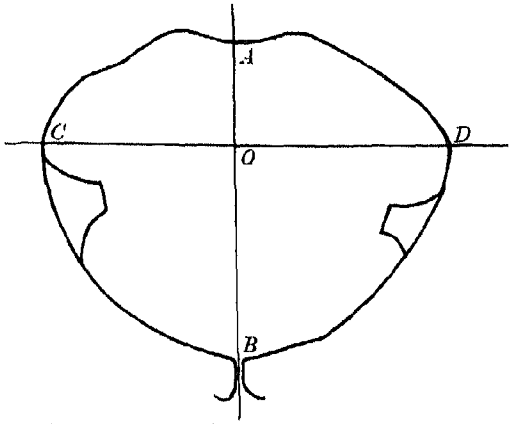


Fig. 5. Outline of the brim of a female pelvis.

Table 2

		Pelvic index				
		76·5	88·5	100·5		
Sagittal index		0 (0·0)	3 (5·3)	20 (16·8)	6 (8·9)	29 (31·0)
	40·5	5 (4·1)	44 (43·3)	76 (72·2)	22 (19·5)	147 (139·1)
	34·5	8 (10·5)	54 (60·1)	55 (53·9)	9 (7·4)	126 (131·9)
	28·5	8 (5·1)	15 (14·6)	4 (6·8)	0 (0·4)	27 (26·9)
		21 (19·7)	116 (123·3)	155 (149·7)	37 (36·2)	329 (328·9)

$r = 0{\cdot}40$, $\kappa/\pi = 0{\cdot}6310$

Table 3

	h	$\frac{h}{\sqrt{(1-r^2)}}$	$\frac{rh}{\sqrt{(1-r^2)}}$	$\frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}dh$
h_1	−1·556	1·698	0·679	0·2201
h_2	−0·165	0·180	0·072	0·0327
h_3	1·226	1·338	0·535	0·1949
k_1	−1·394	1·521	0·608	0·2092
k_2	−0·043	0·047	0·019	0·0086
k_3	1·308	1·427	0·571	0·2028

Table 4

p	h	q_h	$V(h, q_h)$	k	q_k	$V(k, q_k)$	$W(h, k, r)$
1	$-\infty$	—	—	∞	—	—	0·3155
2	$-1·556$	∞	0·2201	∞	—	0·0655	0·2856
3	$-0·165$	∞	0·0327	∞	—	0·0655	0·0982
4	1·226	∞	0·1949	∞	—	$-0·0655$	0·1294
5	∞	—	—	∞	—	—	0·1845
6	$-\infty$	—	0·0655	1·308	∞	0·2028	0·2683
7	$-1·556$	2·106	0·1191	1·308	2·269	0·1193	0·2384
8	$-0·165$	1·499	0·0166	1·308	0·751	0·0505	0·0671
9	1·226	0·892	0·0578	1·308	0·767	0·0515	0·1093
10	∞	—	$-0·0655$	1·308	∞	0·2028	0·1373
11	$-\infty$	—	$-0·0655$	$-0·043$	∞	0·0086	$-0·0569$
12	$-1·556$	$-0·632$	$-0·0442$	$-0·043$	1·679	0·0047	$-0·0395$
13	$-0·165$	$-0·025$	$-0·0003$	$-0·043$	0·161	0·0005	0·0002
14	1·226	0·582	0·0389	$-0·043$	1·357	0·0040	0·0429
15	∞	—	0·0655	$-0·043$	∞	0·0086	0·0741
16	$-\infty$	—	$-0·0655$	$-1·394$	∞	0·2092	0·1437
17	$-1·556$	0·842	0·0577	$-1·394$	1·090	0·0715	0·1292
18	$-0·165$	1·449	0·0161	$-1·394$	$-0·428$	$-0·0300$	$-0·0139$
19	1·226	2·056	0·1099	$-1·394$	1·946	0·1110	0·2209
20	∞	—	0·0655	$-1·394$	∞	0·2092	0·2747
21	$-\infty$	—	—	$-\infty$	—	—	0·1845
22	$-1·556$	∞	0·2201	$-\infty$	—	$-0·0655$	0·1546
23	$-0·165$	∞	0·0327	$-\infty$	—	$-0·0655$	$-0·0328$
24	1·226	∞	0·1949	$-\infty$	—	0·0655	0·2604
25	∞	—	—	$-\infty$	—	—	0·3155

Table 5

Cell	Expected	Observed	χ^2
$a = W(\ 1+\ 7-\ 2-\ 6) = 0·0000$	0·0	0	1·00
$b = W(\ 2+\ 8-\ 3-\ 7) = 0·0161$	5·3	3	
$c = W(\ 3+\ 4-\ 8-\ 9) = 0·0512$	16·8	20	0·61
$d = W(\ 5+\ 9-\ 4-10) = 0·0271$	8·9	6	0·94
$e = W(\ 6+11-\ 7-12) = 0·0125$	4·1	5	0·20
$f = W(\ 7+12-\ 8-13) = 0·1316$	43·3	44	0·01
$g = W(\ 8+13+\ 9+14) = 0·2195$	72·2	76	0·20
$h = W(10+15-\ 9-14) = 0·0592$	19·5	22	0·32
$i = W(12+16-11-17) = 0·0319$	10·5	8	0·60
$j = W(13+17-12-18) = 0·1828$	60·1	54	0·62
$k = W(18+19-13-14) = 0·1639$	53·9	55	0·02
$l = W(14+20-15-19) = 0·0226$	7·4	9	0·35
$m = W(17+21-16-22) = 0·0154$	5·1	8	1·65
$n = W(18+22-17-23) = 0·0443$	14·6	15	0·01
$o = W(23+24-18-19) = 0·0206$	6·8	4	1·42
$p = W(19+25-20-24) = 0·0013$	0·4	0	
1·0000	328·9	329	7·95

Taking the degrees of freedom as 8, this gives a probability of nearly 0·40 of obtaining a worse fit through chance fluctuations, so that we may take it that normal correlation is a good description of this distribution.

Table 6. *Table of $V(h, q)$*

h \ q	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0	q / h
0·1	·000793	·001582	·002364	·003134	·003888	·004625	·005340	·006032	·006699	·007338	0·1
0·2	·001574	·003141	·004692	·006221	·007719	·009182	·010602	·011976	·013300	·014569	0·2
0·3	·002333	·004653	·006952	·009216	·011437	·013604	·015710	·017746	·019708	·021590	0·3
0·4	·003057	·006098	·009110	·012078	·014989	·017830	·020591	·023262	·025836	·028305	0·4
0·5	·003737	·007456	·011139	·014769	·018329	·021805	·025184	·028453	·031604	·034628	0·5
0·6	·004366	·008711	·013014	·017256	·021417	·025481	·029432	·033256	·036943	·040483	0·6
0·7	·004937	·009850	·014716	·019513	·024221	·028819	·033291	·037622	·041799	·045811	0·7
0·8	·005444	·010862	·016229	·021522	·026716	·031792	·036730	·041514	·046130	·050567	0·8
0·9	·005885	·011742	·017545	·023268	·028887	·034380	·039725	·044907	·049909	·054720	0·9
1·0	·006258	·012486	·018659	·024747	·030727	·036574	·042268	·047790	·053124	·058258	1·0
1·1	·006563	·013096	·019571	·025960	·032237	·038378	·044360	·050165	·055777	·061182	1·1
1·2	·006802	·013575	·020288	·026914	·033425	·039799	·046011	·052044	·057880	·063505	1·2
1·3	·006979	·013928	·020817	·027619	·034307	·040855	·047243	·053449	·059458	·065255	1·3
1·4	·007097	·014164	·021172	·028093	·034901	·041571	·048081	·054411	·060544	·066468	1·4
1·5	·007161	·014293	·021366	·028355	·035232	·041974	·048559	·054966	·061180	·067187	1·5
1·6	·007177	·014325	·021418	·028427	·035328	·042097	·048712	·055155	·061409	·067461	1·6
1·7	·007150	·014274	·021342	·028331	·035215	·041972	·048580	·055022	·061281	·067343	1·7
1·8	·007088	·014149	·021159	·028092	·034925	·041635	·048203	·054611	·060843	·066887	1·8
1·9	·006995	·013965	·020885	·027732	·034484	·041120	·047619	·053966	·060145	·066144	1·9
2·0	·006877	·013730	·020537	·027274	·033921	·040458	·046866	·053130	·059234	·065167	2·0
2·1	·006740	·013457	·020130	·026739	·033262	·039682	·045980	·052141	·058153	·064002	2·1
2·2	·006588	·013154	·019680	·026145	·032530	·038818	·044992	·051037	·056941	·062693	2·2
2·3	·006425	·012831	·019199	·025509	·031746	·037891	·043930	·049849	·055636	·061279	2·3
2·4	·006256	·012494	·018697	·024847	·030928	·036924	·042820	·048605	·054267	·059795	2·4
2·5	·006084	·012151	·018185	·024170	·030091	·035933	·041684	·047330	·052862	·058269	2·5
2·6	·005910	·011805	·017670	·023489	·029249	·034936	·040538	·046044	·051442	·056725	2·6
2·7	·005738	·011462	·017158	·022813	·028412	·033944	·039398	·044762	·050027	·055185	2·7
2·8	·005569	·011125	·016655	·022147	·027588	·032968	·038274	·043498	·048630	·053662	2·8
2·9	·005404	·010796	·016165	·021498	·026784	·032013	·037175	·042261	·047262	·052171	2·9
3·0	·005244	·010478	·015689	·020868	·026004	·031087	·036109	·041059	·045932	·050719	3·0

h \ q	1·1	1·2	1·3	1·4	1·5	1·6	1·7	1·8	1·9	2·0	q / h
0·1	·007948	·008529	·009080	·009601	·010092	·010555	·010989	·011395	·011776	·012131	0·1
0·2	·015781	·016935	·018030	·019066	·020042	·020961	·021824	·022633	·023389	·024097	0·2
0·3	·023388	·025100	·026725	·028262	·029712	·031077	·032359	·033560	·034685	·035736	0·3
0·4	·030665	·032913	·035047	·037066	·038973	·040767	·042453	·044035	·045516	·046901	0·4
0·5	·037519	·040274	·042891	·045368	·047708	·049912	·051983	·053927	·055749	·057454	0·5
0·6	·043870	·047098	·050166	·053072	·055818	·058406	·060841	·063127	·065271	·067280	0·6
0·7	·049652	·053315	·056797	·060099	·063221	·066166	·068938	·071544	·073989	·076281	0·7
0·8	·054816	·058872	·062731	·066392	·069856	·073127	·076209	·079108	·081832	·084387	0·8
0·9	·059331	·063735	·067929	·071911	·075683	·079248	·082609	·085775	·088752	·091548	0·9
1·0	·063182	·067889	·072375	·076639	·080682	·084506	·088117	·091521	·094726	·097740	1·0
1·1	·066370	·071334	·076069	·080575	·084852	·088902	·092731	·096345	·099752	·102960	1·1
1·2	·068910	·074086	·079029	·083737	·088212	·092454	·096470	·100266	·103849	·107227	1·2
1·3	·070830	·076175	·081285	·086158	·090795	·095197	·099370	·103320	·107053	·110579	1·3
1·4	·072169	·077642	·082881	·087883	·092648	·097179	·101480	·105557	·109416	·113066	1·4
1·5	·072975	·078537	·083868	·088965	·093828	·098458	·102860	·107039	·111001	·114754	1·5
1·6	·073300	·078917	·084307	·089469	·094400	·099102	·103579	·107836	·111879	·115716	1·6
1·7	·073199	·078840	·084261	·089458	·094431	·099181	·103710	·108024	·112128	·116029	1·7
1·8	·072731	·078369	·083793	·089002	·093993	·098768	·103329	·107680	·111827	·115775	1·8
1·9	·071953	·077563	·082968	·088167	·093156	·097937	·102511	·106882	·111054	·115034	1·9
2·0	·070918	·076481	·081848	·087018	·091987	·096756	·101327	·105703	·109886	·113884	2·0
2·1	·069679	·075178	·080491	·085616	·090550	·095294	·099847	·104214	·108396	·112399	2·1
2·2	·068283	·073704	·078950	·084017	·088904	·093609	·098134	·102479	·106649	·110647	2·2
2·3	·066771	·072104	·077272	·082271	·087100	·091757	·096243	·100558	·104706	·108689	2·3
2·4	·065181	·070417	·075499	·080423	·085186	·089786	·094225	·098502	·102620	·106581	2·4
2·5	·063543	·068678	·073668	·078510	·083201	·087738	·092123	·096355	·100436	·104368	2·5
2·6	·061885	·066914	·071809	·076564	·081178	·085648	·089973	·094155	·098194	·102092	2·6
2·7	·060228	·065150	·069945	·074612	·079145	·083544	·087806	·091934	·095927	·099785	2·7
2·8	·058589	·063402	·068099	·072674	·077125	·081449	·085647	·089717	·093660	·097476	2·8
2·9	·056981	·061687	·066283	·070767	·075134	·079384	·083514	·087524	·091414	·095185	2·9
3·0	·055415	·060014	·064511	·068903	·073187	·077360	·081421	·085370	·089206	·092929	3·0

Table 6 (*continued*)

$h \backslash q$	2·1	2·2	2·3	2·4	2·5	2·6	2·7	2·8	2·9	3·0	∞	q / h
0·1	·012463	·012773	·013062	·013331	·013583	·013818	·014038	·014243	·014435	·014616	·019914	0·1
0·2	·024757	·025374	·025949	·026486	·026987	·027455	·027893	·028302	·028686	·029045	·039630	0·2
0·3	·036719	·037636	·038493	·039292	·040039	·040736	·041389	·041999	·042571	·043107	·058956	0·3
0·4	·048195	·049405	·050535	·051590	·052576	·053497	·054359	·055167	·055923	·056633	·077711	0·4
0·5	·059049	·060539	·061932	·063234	·064451	·065590	·066655	·067654	·068590	·069469	·095731	0·5
0·6	·069159	·070917	·072561	·074099	·075538	·076885	·078147	·079329	·080439	·081482	·112873	0·6
0·7	·078429	·080439	·082321	·084083	·085732	·087277	·088726	·090085	·091361	·092561	·129018	0·7
0·8	·086783	·089029	·091133	·093105	·094953	·096686	·098312	·099839	·101274	·102624	·144072	0·8
0·9	·094173	·096636	·098946	·101113	·103146	·105055	·106848	·108533	·110118	·111612	·157970	0·9
1·0	·100573	·103234	·105733	·108080	·110284	·112356	·114304	·116138	·117865	·119492	·170672	1·0
1·1	·105979	·108819	·111489	·114000	·116362	·118584	·120676	·122648	·124506	·126260	·182167	1·1
1·2	·110411	·113410	·116233	·118893	·121397	·123757	·125981	·128079	·130060	·131932	·192465	1·2
1·3	·113906	·117044	·120004	·122795	·125428	·127912	·130257	·132471	·134564	·136545	·201600	1·3
1·4	·116516	·119776	·122854	·125762	·128508	·131104	·133557	·135877	·138073	·140153	·209622	1·4
1·5	·118308	·121670	·124851	·127860	·130706	·133400	·135950	·138365	·140654	·142825	·216596	1·5
1·6	·119354	·122802	·126069	·129164	·132098	·134878	·137513	·140013	·142386	·144639	·222600	1·6
1·7	·119734	·123252	·126590	·129759	·132766	·135621	·138332	·140907	·143354	·145681	·227717	1·7
1·8	·119531	·123104	·126500	·129729	·132798	·135717	·138492	·141132	·143645	·146038	·232035	1·8
1·9	·118827	·122440	·125882	·129159	·132279	·135251	·138081	·140778	·143348	·145800	·235642	1·9
2·0	·117700	·121343	·124818	·128132	·131293	·134309	·137185	·139930	·142550	·145051	·238625	2·0
2·1	·116227	·119888	·123385	·126727	·129919	·132969	·135884	·138668	·141331	·143876	·241068	2·1
2·2	·114477	·118145	·121656	·125016	·128232	·131308	·134252	·137070	·139767	·142350	·243048	2·2
2·3	·112512	·116179	·119696	·123066	·126296	·129392	·132359	·135203	·137929	·140543	·244638	2·3
2·4	·110389	·114047	·117561	·120935	·124173	·127282	·130265	·133129	·135878	·138517	·245901	2·4
2·5	·108154	·111798	·115303	·118674	·121915	·125029	·128024	·130902	·133668	·136328	·246895	2·5
2·6	·105851	·109474	·112965	·116327	·119564	·122680	·125680	·128567	·131346	·134021	·247669	2·6
2·7	·103513	·107111	·110582	·113931	·117160	·120272	·123273	·126164	·128951	·131637	·248267	2·7
2·8	·101168	·104736	·108185	·111516	·114732	·117837	·120834	·123726	·126517	·129210	·248722	2·8
2·9	·098838	·102374	·105796	·109106	·112306	·115399	·118389	·121278	·124069	·126766	·249067	2·9
3·0	·096541	·100042	·103435	·106720	·109901	·112980	·115959	·118841	·121629	·124326	·249325	3·0

$$V(0, q)=0, \quad V(h, 0)=0, \quad V(\infty, q)=\frac{1}{2\pi}\tan^{-1}\left\{\frac{-r}{\sqrt{(1-r^2)}}\right\}, \quad V(h, \infty)=\frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}\,dh,$$

$$W(\infty, \infty, r)=\frac{1}{\pi}\tan^{-1}\left\{\frac{\sqrt{(1-r)}}{\sqrt{(1+r)}}\right\}, \quad \text{for } q>3 \; V(h, q)=\frac{1}{2}\frac{1}{\sqrt{(2\pi)}}\int_0^h e^{-\frac{1}{2}h^2}\,dh-\frac{1}{2\pi}\tan^{-1}(h/q).$$

Table 7. *Auxiliary table of trigonometric functions*

r	$\sqrt{(1-r^2)}$	$1/\sqrt{(1-r^2)}$	$r/\sqrt{(1-r^2)}$	$\kappa/\pi-\frac{1}{2}$	r	$\sqrt{(1-r^2)}$	$1/\sqrt{(1-r^2)}$	$r/\sqrt{(1-r^2)}$	$\kappa/\pi-\frac{1}{2}$
$\sin\lambda$	$\cos\lambda$	$\sec\lambda$	$\tan\lambda$	λ/π	$\sin\lambda$	$\cos\lambda$	$\sec\lambda$	$\tan\lambda$	λ/π
0·00	1·00000	1·00000	0·00000	0·000000	**0·50**	0·86603	1·15470	0·57735	0·166667
0·01	0·99995	1·00005	0·01000	0·003183	**0·51**	0·86017	1·16255	0·59290	0·170355
·02	·99980	1·00020	·02000	·006367	**·52**	·85417	1·17073	·60878	·174068
·03	·99955	1·00045	·03001	·009551	**·53**	·84800	1·17925	·62500	·177808
·04	·99920	1·00080	·04003	·012736	**·54**	·84167	1·18812	·64159	·181576
·05	·99875	1·00125	·05006	·015922	**·55**	·83516	1·19737	·65855	·185373
0·06	0·99820	1·00180	0·06011	0·019110	**0·56**	0·82849	1·20701	0·67593	0·189199
·07	·99755	1·00246	·07017	·022300	**·57**	·82164	1·21707	·69373	·193057
·08	·99679	1·00322	·08026	·025492	**·58**	·81462	1·22757	·71199	·196948
·09	·99599	1·00407	·09037	·028687	**·59**	·80740	1·23854	·73074	·200872
·10	·99499	1·00504	·10050	·031884	**·60**	·80000	1·25000	·75000	·204833
0·11	0·99393	1·00611	0·11067	0·035085	**0·61**	0·79240	1·26199	0·76981	0·208831
·12	·99277	1·00728	·12087	·038289	**·62**	·78460	1·27453	·79021	·212867
·13	·99151	1·00856	·13111	·041498	**·63**	·77660	1·28767	·81123	·216945
·14	·99015	1·00995	·14139	·044710	**·64**	·76837	1·30145	·83293	·221066
·15	·98869	1·01144	·15172	·047927	**·65**	·75993	1·31590	·85534	·225231
0·16	0·98712	1·01305	0·16209	0·051150	**0·66**	0·75127	1·33109	0·87852	0·229443
·17	·98544	1·01477	·17251	·054377	**·67**	·74236	1·34705	·90253	·233706
·18	·98367	1·01660	·18299	·057610	**·68**	·73321	1·36386	·92743	·238020
·19	·98178	1·01855	·19353	·060849	**·69**	·72381	1·38158	·95329	·242390
·20	·97980	1·02062	·20412	·064094	**·70**	·71414	1·40028	·98020	·246817
0·21	0·97770	1·02281	0·21479	0·067346	**0·71**	0·70420	1·42005	1·00823	0·251305
·22	·97550	1·02512	·22553	·070606	**·72**	·69397	1·44098	1·03750	·255858
·23	·97319	1·02755	·23634	·073873	**·73**	·68345	1·46317	1·06811	·260480
·24	·97077	1·03011	·24723	·077147	**·74**	·67261	1·48675	1·10020	·265174
·25	·96825	1·03280	·25820	·080431	**·75**	·66144	1·51186	1·13389	·269946
0·26	0·96561	1·03562	0·26926	0·083723	**0·76**	0·64992	1·53864	1·16937	0·274801
·27	·96280	1·03857	·28041	·087024	**·77**	·63804	1·56729	1·20681	·279744
·28	·96000	1·04167	·29167	·090335	**·78**	·62578	1·59801	1·24645	·284781
·29	·95703	1·04490	·30302	·093655	**·79**	·61311	1·63104	1·28852	·289919
·30	·95394	1·04828	·31449	·096987	**·80**	·60000	1·66667	1·33333	·295167
0·31	0·95074	1·05182	0·32606	0·100329	**0·81**	0·58643	1·70523	1·38124	0·300533
·32	·94742	1·05550	·33776	·103683	**·82**	·57236	1·74714	1·43266	·306027
·33	·94398	1·05934	·34958	·107049	**·83**	·55776	1·79287	1·48809	·311660
·34	·94043	1·06335	·36154	·110427	**·84**	·54259	1·84302	1·54814	·317445
·35	·93675	1·06752	·37363	·113818	**·85**	·52678	1·89832	1·61357	·323398
0·36	0·93295	1·07187	0·38587	0·117223	**0·86**	0·51029	1·95965	1·68530	0·329537
·37	·92903	1·07639	·39826	·120642	**·87**	·49305	2·02818	1·76452	·335881
·38	·92499	1·08110	·41082	·124076	**·88**	·47497	2·10358	1·85273	·342458
·39	·92081	1·08599	·42354	·127525	**·89**	·45596	2·19317	1·95192	·349296
·40	·91652	1·09109	·43644	·130990	**·90**	·43589	2·29416	2·06474	·356433
0·41	0·91209	1·09639	0·44952	0·134471	**0·91**	0·41461	2·41192	2·19484	0·363919
·42	·90752	1·10190	·46280	·137970	**·92**	·39192	2·55155	2·34743	·371812
·43	·90283	1·10763	·47628	·141487	**·93**	·36756	2·72065	2·53020	·380193
·44	·89800	1·11359	·48998	·145022	**·94**	·34117	2·93105	2·75519	·389175
·45	·89303	1·11979	·50390	·148576	**·95**	·31225	3·20256	3·04243	·398917
0·46	0·88792	1·12623	0·51807	0·152151	**0·96**	0·28000	3·57143	3·42857	0·409666
·47	·88267	1·13293	·53248	·155746	**·97**	·24310	4·11343	3·99003	·421834
·48	·87727	1·13990	·54715	·159363	**·98**	·19900	5·02519	4·92469	·436231
·49	·87172	1·14715	·56211	·163003	**·99**	·14107	7·08881	7·01792	·454947
·50	·86603	1·15470	·57735	·166667	**1·00**	·00000	∞	∞	·500000

REFERENCES

Caldwell, W. E. & Moloy, H. C. (1938). *Proc. Roy. Soc. Med.* **32**, 1.

Elderton, Ethel M., Moul, Margaret, Fieller, E. C., Pretorius, S. J. & Church, A. E. R. (1930). *Biometrika*, **22**, 1.

Everitt, P. F. (1910). *Biometrika*, **7**, 437.

Everitt, P. F. (1912). *Biometrika*, **8**, 385.

Lee, Alice (1915). *Biometrika*, **11**, 284.

Lee, Alice (1927). *Biometrika*, **19**, 354.

Nicholson, C. (1936). *Lancet*, no. 231, p. 615.

Nicholson, C. (1941). *Biometrika*, **32**, 16.

Pearson, K. (1901). *Philos. Trans.* A, **195**, 1.

Sheppard, W. F. (1900). *Trans. Camb. Phil. Soc.* **19**, 23.

Turner, Sir W. (1886). *Report on the Scientific Results of the Voyage of H.M.S. 'Challenger'*, **16**, 47. London.

TABLES OF PERCENTAGE POINTS OF THE INVERTED BETA (F) DISTRIBUTION

COMPUTED BY MAXINE MERRINGTON AND CATHERINE M. THOMPSON

PREFATORY NOTE BY E. S. PEARSON

The following tables of the percentage points of F, using the notation of Snedecor (1934), express the results of Miss Thompson's tabulation (1941 a) of the incomplete beta function in terms of the argument most convenient for use in the analysis of variance.

If we take the elementary probability function* of the beta distribution, namely

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)}\, x^{p-1}(1-x)^{q-1} \tag{1}$$

and make the transformation

$$x = \frac{1}{1+u}, \tag{2}$$

we obtain for u the inverted beta distribution

$$f(u) = \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)}\, u^{q-1}(1+u)^{-p-q}. \tag{3}$$

The limits 0 and 1 for x correspond to limits of ∞ and 0 for u. While equation (1) represents a standardized form of Karl Pearson's Type I frequency distribution, (3) is a form of his Type VI. The *Tables of Incomplete Beta-Function* (K. Pearson 1934) provide the probability integral of (1) and therefore of (3).

In the terminology of the analysis of variance let S_1 and S_2 be two sums of squares of normal variates having, respectively, ν_1 and ν_2 degrees of freedom. If all the variates have a common standard deviation σ and if S_1 and S_2 are independent, then it is known that:

(a) S_1/σ^2 and S_2/σ^2 are distributed in the standard χ^2 form, namely

$$f(\chi^2) = \frac{(\frac{1}{2})^{\frac{1}{2}\nu}}{\Gamma(\frac{1}{2}\nu)}\,(\chi^2)^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}\chi^2} \tag{4}$$

with ν_1 and ν_2 degrees of freedom, respectively.

(b) The ratio S_1/S_2 is distributed as u in (3), where $q = \frac{1}{2}\nu_1$, $p = \frac{1}{2}\nu_2$.

(c) Writing

$$s_1^2 = \frac{S_1}{\nu_1}, \quad s_2^2 = \frac{S_2}{\nu_2} \tag{5}$$

as two independent estimates of σ^2, the ratio

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2} \tag{6}$$

has a probability distribution

$$f(F) = \frac{\Gamma(\frac{1}{2}\nu_1 + \frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\, \nu_1^{\frac{1}{2}\nu_1}\nu_2^{\frac{1}{2}\nu_2} F^{\frac{1}{2}\nu_1-1}(\nu_2+\nu_1 F)^{-\frac{1}{2}(\nu_1+\nu_2)}. \tag{7}$$

It is useful to note that, for (7),

$$\text{Expectation of } F = \frac{\nu_2}{\nu_2-2}, \qquad \text{for } \nu_2 > 2, \tag{8}$$

$$\sigma_F = \frac{\nu_2}{\nu_2-2}\sqrt{\left\{\frac{2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}\right\}}, \qquad \text{for } \nu_2 > 4. \tag{9}$$

Thus for large values of ν_2, F tends to be distributed as χ^2/ν_1 with a mean of unity and standard deviation of $\sqrt{(2/\nu_1)}$.

* The letter f will be used as a general symbol for an elementary probability function, in place of p which might here be confused with the index. The integral probability function for the beta distribution is then

$$P\{0 \leqslant x \leqslant X\} = \int_0^X f(x)\,dx.$$

F Distribution: 50 per cent Points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	1·0000	1·5000	1·7092	1·8227	1·8937	1·9422	1·9774	2·0041	2·0250
2	0·66667	1·0000	1·1349	1·2071	1·2519	1·2824	1·3045	1·3213	1·3344
3	·58506	0·88110	1·0000	1·0632	1·1024	1·1289	1·1482	1·1627	1·1741
4	·54863	·82843	0·94054	1·0000	1·0367	1·0617	1·0797	1·0933	1·1040
5	0·52807	0·79877	0·90715	0·96456	1·0000	1·0240	1·0414	1·0545	1·0648
6	·51489	·77976	·88578	·94191	0·97654	1·0000	1·0169	1·0298	1·0398
7	·50572	·76655	·87095	·92619	·96026	0·98334	1·0000	1·0126	1·0224
8	·49898	·75683	·86004	·91464	·94831	·97111	0·98757	1·0000	1·0097
9	·49382	·74938	·85168	·90580	·93916	·96175	·97805	0·99037	1·0000
10	0·48973	0·74349	0·84508	0·89882	0·93193	0·95436	0·97054	0·98276	0·99232
11	·48644	·73872	·83973	·89316	·92608	·94837	·96445	·97661	·98610
12	·48369	·73477	·83530	·88848	·92124	·94342	·95943	·97152	·98097
13	·48141	·73145	·83159	·88454	·91718	·93926	·95520	·96724	·97665
14	·47944	·72862	·82842	·88119	·91371	·93573	·95161	·96360	·97298
15	0·47775	0·72619	0·82569	0·87830	0·91073	0·93267	0·94850	0·96046	0·96981
16	·47628	·72406	·82330	·87578	·90812	·93001	·94580	·95773	·96705
17	·47499	·72219	·82121	·87357	·90584	·92767	·94342	·95532	·96462
18	·47385	·72053	·81936	·87161	·90381	·92560	·94132	·95319	·96247
19	·47284	·71906	·81771	·86987	·90200	·92375	·93944	·95129	·96056
20	0·47192	0·71773	0·81621	0·86830	0·90038	0·92210	0·93776	0·94959	0·95884
21	·47108	·71653	·81487	·86688	·89891	·92060	·93624	·94805	·95728
22	·47033	·71545	·81365	·86559	·89759	·91924	·93486	·94665	·95588
23	·46965	·71446	·81255	·86442	·89638	·91800	·93360	·94538	·95459
24	·46902	·71356	·81153	·86335	·89527	·91687	·93245	·94422	·95342
25	0·46844	0·71272	0·81061	0·86236	0·89425	0·91583	0·93140	0·94315	0·95234
26	·46793	·71195	·80975	·86145	·89331	·91487	·93042	·94217	·95135
27	·46744	·71124	·80894	·86061	·89244	·91399	·92952	·94126	·95044
28	·46697	·71059	·80820	·85983	·89164	·91317	·92869	·94041	·94958
29	·46654	·70999	·80753	·85911	·89089	·91241	·92791	·93963	·94879
30	0·46616	0·70941	0·80689	0·85844	0·89019	0·91169	0·92719	0·93889	0·94805
40	·46330	·70531	·80228	·85357	·88516	·90654	·92197	·93361	·94272
60	·46053	·70122	·79770	·84873	·88017	·90144	·91679	·92838	·93743
120	·45774	·69717	·79314	·84392	·87521	·89637	·91164	·92318	·93218
∞	·45494	·69315	·78866	·83918	·87029	·89135	·90654	·91802	·92698

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}50$.

F DISTRIBUTION: 50 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	2·0419	2·0674	2·0931	2·1190	2·1321	2·1452	2·1584	2·1716	2·1848	2·1981
2	1·3450	1·3610	1·3771	1·3933	1·4014	1·4096	1·4178	1·4261	1·4344	1·4427
3	1·1833	1·1972	1·2111	1·2252	1·2322	1·2393	1·2464	1·2536	1·2608	1·2680
4	1·1126	1·1255	1·1386	1·1517	1·1583	1·1649	1·1716	1·1782	1·1849	1·1916
5	1·0730	1·0855	1·0980	1·1106	1·1170	1·1234	1·1297	1·1361	1·1426	1·1490
6	1·0478	1·0600	1·0722	1·0845	1·0907	1·0969	1·1031	1·1093	1·1156	1·1219
7	1·0304	1·0423	1·0543	1·0664	1·0724	1·0785	1·0846	1·0908	1·0969	1·1031
8	1·0175	1·0293	1·0412	1·0531	1·0591	1·0651	1·0711	1·0771	1·0832	1·0893
9	1·0077	1·0194	1·0311	1·0429	1·0489	1·0548	1·0608	1·0667	1·0727	1·0788
10	1·0000	1·0116	1·0232	1·0349	1·0408	1·0467	1·0526	1·0585	1·0645	1·0705
11	0·99373	1·0052	1·0168	1·0284	1·0343	1·0401	1·0460	1·0519	1·0578	1·0637
12	·98856	1·0000	1·0115	1·0231	1·0289	1·0347	1·0405	1·0464	1·0523	1·0582
13	·98421	0·99560	1·0071	1·0186	1·0243	1·0301	1·0360	1·0418	1·0476	1·0535
14	·98051	·99186	1·0033	1·0147	1·0205	1·0263	1·0321	1·0379	1·0437	1·0495
15	0·97732	0·98863	1·0000	1·0114	1·0172	1·0229	1·0287	1·0345	1·0403	1·0461
16	·97454	·98582	0·99716	1·0086	1·0143	1·0200	1·0258	1·0315	1·0373	1·0431
17	·97209	·98334	·99466	1·0060	1·0117	1·0174	1·0232	1·0289	1·0347	1·0405
18	·96993	·98116	·99245	1·0038	1·0095	1·0152	1·0209	1·0267	1·0324	1·0382
19	·96800	·97920	·99047	1·0018	1·0075	1·0132	1·0189	1·0246	1·0304	1·0361
20	0·96626	0·97746	0·98870	1·0000	1·0057	1·0114	1·0171	1·0228	1·0285	1·0343
21	·96470	·97587	·98710	0·99838	1·0040	1·0097	1·0154	1·0211	1·0268	1·0326
22	·96328	·97444	·98565	·99692	1·0026	1·0082	1·0139	1·0196	1·0253	1·0311
23	·96199	·97313	·98433	·99558	1·0012	1·0069	1·0126	1·0183	1·0240	1·0297
24	·96081	·97194	·98312	·99436	1·0000	1·0057	1·0113	1·0170	1·0227	1·0284
25	0·95972	0·97084	0·98201	0·99324	0·99887	1·0045	1·0102	1·0159	1·0215	1·0273
26	·95872	·96983	·98099	·99220	·99783	1·0035	1·0091	1·0148	1·0205	1·0262
27	·95779	·96889	·98004	·99125	·99687	1·0025	1·0082	1·0138	1·0195	1·0252
28	·95694	·96802	·97917	·99036	·99598	1·0016	1·0073	1·0129	1·0186	1·0243
29	·95614	·96722	·97835	·98954	·99515	1·0008	1·0064	1·0121	1·0177	1·0234
30	0·95540	0·96647	0·97759	0·98877	0·99438	1·0000	1·0056	1·0113	1·0170	1·0226
40	·95003	·96104	·97211	·98323	·98880	0·99440	1·0000	1·0056	1·0113	1·0169
60	·94471	·95566	·96667	·97773	·98328	·98884	0·99441	1·0000	1·0056	1·0112
120	·93943	·95032	·96128	·97228	·97780	·98333	·98887	0·99443	1·0000	1·0056
∞	·93418	·94503	·95593	·96687	·97236	·97787	·98339	·98891	0·99445	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

F Distribution: 25 per cent Points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	5·8285	7·5000	8·1999	8·5810	8·8198	8·9833	9·1021	9·1922	9·2631
2	2·5714	3·0000	3·1534	3·2320	3·2799	3·3121	3·3352	3·3526	3·3661
3	2·0239	2·2798	2·3555	2·3901	2·4095	2·4218	2·4302	2·4364	2·4410
4	1·8074	2·0000	2·0467	2·0642	2·0723	2·0766	2·0790	2·0805	2·0814
5	1·6925	1·8528	1·8843	1·8927	1·8947	1·8945	1·8935	1·8923	1·8911
6	1·6214	1·7622	1·7844	1·7872	1·7852	1·7821	1·7789	1·7760	1·7733
7	1·5732	1·7010	1·7169	1·7157	1·7111	1·7059	1·7011	1·6969	1·6931
8	1·5384	1·6569	1·6683	1·6642	1·6575	1·6508	1·6448	1·6396	1·6350
9	1·5121	1·6236	1·6315	1·6253	1·6170	1·6091	1·6022	1·5961	1·5909
10	1·4915	1·5975	1·6028	1·5949	1·5853	1·5765	1·5688	1·5621	1·5563
11	1·4749	1·5767	1·5798	1·5704	1·5598	1·5502	1·5418	1·5346	1·5284
12	1·4613	1·5595	1·5609	1·5503	1·5389	1·5286	1·5197	1·5120	1·5054
13	1·4500	1·5452	1·5451	1·5336	1·5214	1·5105	1·5011	1·4931	1·4861
14	1·4403	1·5331	1·5317	1·5194	1·5066	1·4952	1·4854	1·4770	1·4697
15	1·4321	1·5227	1·5202	1·5071	1·4938	1·4820	1·4718	1·4631	1·4556
16	1·4249	1·5137	1·5103	1·4965	1·4827	1·4705	1·4601	1·4511	1·4433
17	1·4186	1·5057	1·5015	1·4873	1·4730	1·4605	1·4497	1·4405	1·4325
18	1·4130	1·4988	1·4938	1·4790	1·4644	1·4516	1·4406	1·4312	1·4230
19	1·4081	1·4925	1·4870	1·4717	1·4568	1·4437	1·4325	1·4228	1·4145
20	1·4037	1·4870	1·4808	1·4652	1·4500	1·4366	1·4252	1·4153	1·4069
21	1·3997	1·4820	1·4753	1·4593	1·4438	1·4302	1·4186	1·4086	1·4000
22	1·3961	1·4774	1·4703	1·4540	1·4382	1·4244	1·4126	1·4025	1·3937
23	1·3928	1·4733	1·4657	1·4491	1·4331	1·4191	1·4072	1·3969	1·3880
24	1·3898	1·4695	1·4615	1·4447	1·4285	1·4143	1·4022	1·3918	1·3828
25	1·3870	1·4661	1·4577	1·4406	1·4242	1·4099	1·3976	1·3871	1·3780
26	1·3845	1·4629	1·4542	1·4368	1·4203	1·4058	1·3935	1·3828	1·3737
27	1·3822	1·4600	1·4510	1·4334	1·4166	1·4021	1·3896	1·3788	1·3696
28	1·3800	1·4572	1·4480	1·4302	1·4133	1·3986	1·3860	1·3752	1·3658
29	1·3780	1·4547	1·4452	1·4272	1·4102	1·3953	1·3826	1·3717	1·3623
30	1·3761	1·4524	1·4426	1·4244	1·4073	1·3923	1·3795	1·3685	1·3590
40	1·3626	1·4355	1·4239	1·4045	1·3863	1·3706	1·3571	1·3455	1·3354
60	1·3493	1·4188	1·4055	1·3848	1·3657	1·3491	1·3349	1·3226	1·3119
120	1·3362	1·4024	1·3873	1·3654	1·3453	1·3278	1·3128	1·2999	1·2886
∞	1·3233	1·3863	1·3694	1·3463	1·3251	1·3068	1·2910	1·2774	1·2654

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}25$.

F DISTRIBUTION: 25 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	9·3202	9·4064	9·4934	9·5813	9·6255	9·6698	9·7144	9·7591	9·8041	9·8492
2	3·3770	3·3934	3·4098	3·4263	3·4345	3·4428	3·4511	3·4594	3·4677	3·4761
3	2·4447	2·4500	2·4552	2·4602	2·4626	2·4650	2·4674	2·4697	2·4720	2·4742
4	2·0820	2·0826	2·0829	2·0828	2·0827	2·0825	2·0821	2·0817	2·0812	2·0806
5	1·8899	1·8877	1·8851	1·8820	1·8802	1·8784	1·8763	1·8742	1·8719	1·8694
6	1·7708	1·7668	1·7621	1·7569	1·7540	1·7510	1·7477	1·7443	1·7407	1·7368
7	1·6898	1·6843	1·6781	1·6712	1·6675	1·6635	1·6593	1·6548	1·6502	1·6452
8	1·6310	1·6244	1·6170	1·6088	1·6043	1·5996	1·5945	1·5892	1·5836	1·5777
9	1·5863	1·5788	1·5705	1·5611	1·5560	1·5506	1·5450	1·5389	1·5325	1·5257
10	1·5513	1·5430	1·5338	1·5235	1·5179	1·5119	1·5056	1·4990	1·4919	1·4843
11	1·5230	1·5140	1·5041	1·4930	1·4869	1·4805	1·4737	1·4664	1·4587	1·4504
12	1·4996	1·4902	1·4796	1·4678	1·4613	1·4544	1·4471	1·4393	1·4310	1·4221
13	1·4801	1·4701	1·4590	1·4465	1·4397	1·4324	1·4247	1·4164	1·4075	1·3980
14	1·4634	1·4530	1·4414	1·4284	1·4212	1·4136	1·4055	1·3967	1·3874	1·3772
15	1·4491	1·4383	1·4263	1·4127	1·4052	1·3973	1·3888	1·3796	1·3698	1·3591
16	1·4366	1·4255	1·4130	1·3990	1·3913	1·3830	1·3742	1·3646	1·3543	1·3432
17	1·4256	1·4142	1·4014	1·3869	1·3790	1·3704	1·3613	1·3514	1·3406	1·3290
18	1·4159	1·4042	1·3911	1·3762	1·3680	1·3592	1·3497	1·3395	1·3284	1·3162
19	1·4073	1·3953	1·3819	1·3666	1·3582	1·3492	1·3394	1·3289	1·3174	1·3048
20	1·3995	1·3873	1·3736	1·3580	1·3494	1·3401	1·3301	1·3193	1·3074	1·2943
21	1·3925	1·3801	1·3661	1·3502	1·3414	1·3319	1·3217	1·3105	1·2983	1·2848
22	1·3861	1·3735	1·3593	1·3431	1·3341	1·3245	1·3140	1·3025	1·2900	1·2761
23	1·3803	1·3675	1·3531	1·3366	1·3275	1·3176	1·3069	1·2952	1·2824	1·2681
24	1·3750	1·3621	1·3474	1·3307	1·3214	1·3113	1·3004	1·2885	1·2754	1·2607
25	1·3701	1·3570	1·3422	1·3252	1·3158	1·3056	1·2945	1·2823	1·2689	1·2538
26	1·3656	1·3524	1·3374	1·3202	1·3106	1·3002	1·2889	1·2765	1·2628	1·2474
27	1·3615	1·3481	1·3329	1·3155	1·3058	1·2953	1·2838	1·2712	1·2572	1·2414
28	1·3576	1·3441	1·3288	1·3112	1·3013	1·2906	1·2790	1·2662	1·2519	1·2358
29	1·3541	1·3404	1·3249	1·3071	1·2971	1·2863	1·2745	1·2615	1·2470	1·2306
30	1·3507	1·3369	1·3213	1·3033	1·2933	1·2823	1·2703	1·2571	1·2424	1·2256
40	1·3266	1·3119	1·2952	1·2758	1·2649	1·2529	1·2397	1·2249	1·2080	1·1883
60	1·3026	1·2870	1·2691	1·2481	1·2361	1·2229	1·2081	1·1912	1·1715	1·1474
120	1·2787	1·2621	1·2428	1·2200	1·2068	1·1921	1·1752	1·1555	1·1314	1·0987
∞	1·2549	1·2371	1·2163	1·1914	1·1767	1·1600	1·1404	1·1164	1·0838	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

F Distribution: 10 per cent Points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	39·864	49·500	53·593	55·833	57·241	58·204	58·906	59·439	59·858
2	8·5263	9·0000	9·1618	9·2434	9·2926	9·3255	9·3491	9·3668	9·3805
3	5·5383	5·4624	5·3908	5·3427	5·3092	5·2847	5·2662	5·2517	5·2400
4	4·5448	4·3246	4·1908	4·1073	4·0506	4·0098	3·9790	3·9549	3·9357
5	4·0604	3·7797	3·6195	3·5202	3·4530	3·4045	3·3679	3·3393	3·3163
6	3·7760	3·4633	3·2888	3·1808	3·1075	3·0546	3·0145	2·9830	2·9577
7	3·5894	3·2574	3·0741	2·9605	2·8833	2·8274	2·7849	2·7516	2·7247
8	3·4579	3·1131	2·9238	2·8064	2·7265	2·6683	2·6241	2·5893	2·5612
9	3·3603	3·0065	2·8129	2·6927	2·6106	2·5509	2·5053	2·4694	2·4403
10	3·2850	2·9245	2·7277	2·6053	2·5216	2·4606	2·4140	2·3772	2·3473
11	3·2252	2·8595	2·6602	2·5362	2·4512	2·3891	2·3416	2·3040	2·2735
12	3·1765	2·8068	2·6055	2·4801	2·3940	2·3310	2·2828	2·2446	2·2135
13	3·1362	2·7632	2·5603	2·4337	2·3467	2·2830	2·2341	2·1953	2·1638
14	3·1022	2·7265	2·5222	2·3947	2·3069	2·2426	2·1931	2·1539	2·1220
15	3·0732	2·6952	2·4898	2·3614	2·2730	2·2081	2·1582	2·1185	2·0862
16	3·0481	2·6682	2·4618	2·3327	2·2438	2·1783	2·1280	2·0880	2·0553
17	3·0262	2·6446	2·4374	2·3077	2·2183	2·1524	2·1017	2·0613	2·0284
18	3·0070	2·6239	2·4160	2·2858	2·1958	2·1296	2·0785	2·0379	2·0047
19	2·9899	2·6056	2·3970	2·2663	2·1760	2·1094	2·0580	2·0171	1·9836
20	2·9747	2·5893	2·3801	2·2489	2·1582	2·0913	2·0397	1·9985	1·9649
21	2·9609	2·5746	2·3649	2·2333	2·1423	2·0751	2·0232	1·9819	1·9480
22	2·9486	2·5613	2·3512	2·2193	2·1279	2·0605	2·0084	1·9668	1·9327
23	2·9374	2·5493	2·3387	2·2065	2·1149	2·0472	1·9949	1·9531	1·9189
24	2·9271	2·5383	2·3274	2·1949	2·1030	2·0351	1·9826	1·9407	1·9063
25	2·9177	2·5283	2·3170	2·1843	2·0922	2·0241	1·9714	1·9292	1·8947
26	2·9091	2·5191	2·3075	2·1745	2·0822	2·0139	1·9610	1·9188	1·8841
27	2·9012	2·5106	2·2987	2·1655	2·0730	2·0045	1·9515	1·9091	1·8743
28	2·8939	2·5028	2·2906	2·1571	2·0645	1·9959	1·9427	1·9001	1·8652
29	2·8871	2·4955	2·2831	2·1494	2·0566	1·9878	1·9345	1·8918	1·8568
30	2·8807	2·4887	2·2761	2·1422	2·0492	1·9803	1·9269	1·8841	1·8490
40	2·8354	2·4404	2·2261	2·0909	1·9968	1·9269	1·8725	1·8289	1·7929
60	2·7914	2·3932	2·1774	2·0410	1·9457	1·8747	1·8194	1·7748	1·7380
120	2·7478	2·3473	2·1300	1·9923	1·8959	1·8238	1·7675	1·7220	1·6843
∞	2·7055	2·3026	2·0838	1·9449	1·8473	1·7741	1·7167	1·6702	1·6315

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}10$.

F DISTRIBUTION: 10 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	60·195	60·705	61·220	61·740	62·002	62·265	62·529	62·794	63·061	63·328
2	9·3916	9·4081	9·4247	9·4413	9·4496	9·4579	9·4663	9·4746	9·4829	9·4913
3	5·2304	5·2156	5·2003	5·1845	5·1764	5·1681	5·1597	5·1512	5·1425	5·1337
4	3·9199	3·8955	3·8689	3·8443	3·8310	3·8174	3·8036	3·7896	3·7753	3·7607
5	3·2974	3·2682	3·2380	3·2067	3·1905	3·1741	3·1573	3·1402	3·1228	3·1050
6	2·9369	2·9047	2·8712	2·8363	2·8183	2·8000	2·7812	2·7620	2·7423	2·7222
7	2·7025	2·6681	2·6322	2·5947	2·5753	2·5555	2·5351	2·5142	2·4928	2·4708
8	2·5380	2·5020	2·4642	2·4246	2·4041	2·3830	2·3614	2·3391	2·3162	2·2926
9	2·4163	2·3789	2·3396	2·2983	2·2768	2·2547	2·2320	2·2085	2·1843	2·1592
10	2·3226	2·2841	2·2435	2·2007	2·1784	2·1554	2·1317	2·1072	2·0818	2·0554
11	2·2482	2·2087	2·1671	2·1230	2·1000	2·0762	2·0516	2·0261	1·9997	1·9721
12	2·1878	2·1474	2·1049	2·0597	2·0360	2·0115	1·9861	1·9597	1·9323	1·9036
13	2·1376	2·0966	2·0532	2·0070	1·9827	1·9576	1·9315	1·9043	1·8759	1·8462
14	2·0954	2·0537	2·0095	1·9625	1·9377	1·9119	1·8852	1·8572	1·8280	1·7973
15	2·0593	2·0171	1·9722	1·9243	1·8990	1·8728	1·8454	1·8168	1·7867	1·7551
16	2·0281	1·9854	1·9399	1·8913	1·8656	1·8388	1·8108	1·7816	1·7507	1·7182
17	2·0009	1·9577	1·9117	1·8624	1·8362	1·8090	1·7805	1·7506	1·7191	1·6856
18	1·9770	1·9333	1·8868	1·8368	1·8103	1·7827	1·7537	1·7232	1·6910	1·6567
19	1·9557	1·9117	1·8647	1·8142	1·7873	1·7592	1·7298	1·6988	1·6659	1·6308
20	1·9367	1·8924	1·8449	1·7938	1·7667	1·7382	1·7083	1·6768	1·6433	1·6074
21	1·9197	1·8750	1·8272	1·7756	1·7481	1·7193	1·6890	1·6569	1·6228	1·5862
22	1·9043	1·8593	1·8111	1·7590	1·7312	1·7021	1·6714	1·6389	1·6042	1·5668
23	1·8903	1·8450	1·7964	1·7439	1·7159	1·6864	1·6554	1·6224	1·5871	1·5490
24	1·8775	1·8319	1·7831	1·7302	1·7019	1·6721	1·6407	1·6073	1·5715	1·5327
25	1·8658	1·8200	1·7708	1·7175	1·6890	1·6589	1·6272	1·5934	1·5570	1·5176
26	1·8550	1·8090	1·7596	1·7059	1·6771	1·6468	1·6147	1·5805	1·5437	1·5036
27	1·8451	1·7989	1·7492	1·6951	1·6662	1·6356	1·6032	1·5686	1·5313	1·4906
28	1·8359	1·7895	1·7395	1·6852	1·6560	1·6252	1·5925	1·5575	1·5198	1·4784
29	1·8274	1·7808	1·7306	1·6759	1·6465	1·6155	1·5825	1·5472	1·5090	1·4670
30	1·8195	1·7727	1·7223	1·6673	1·6377	1·6065	1·5732	1·5376	1·4989	1·4564
40	1·7627	1·7146	1·6624	1·6052	1·5741	1·5411	1·5056	1·4672	1·4248	1·3769
60	1·7070	1·6574	1·6034	1·5435	1·5107	1·4755	1·4373	1·3952	1·3476	1·2915
120	1·6524	1·6012	1·5450	1·4821	1·4472	1·4094	1·3676	1·3203	1·2646	1·1926
∞	1·5987	1·5458	1·4871	1·4206	1·3832	1·3419	1·2951	1·2400	1·1686	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

F Distribution: 5 per cent Points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161·45	199·50	215·71	224·58	230·16	233·99	236·77	238·88	240·54
2	18·513	19·000	19·164	19·247	19·296	19·330	19·353	19·371	19·385
3	10·128	9·5521	9·2766	9·1172	9·0135	8·9406	8·8868	8·8452	8·8123
4	7·7086	6·9443	6·5914	6·3883	6·2560	6·1631	6·0942	6·0410	5·9988
5	6·6079	5·7861	5·4095	5·1922	5·0503	4·9503	4·8759	4·8183	4·7725
6	5·9874	5·1433	4·7571	4·5337	4·3874	4·2839	4·2066	4·1468	4·0990
7	5·5914	4·7374	4·3468	4·1203	3·9715	3·8660	3·7870	3·7257	3·6767
8	5·3177	4·4590	4·0662	3·8378	3·6875	3·5806	3·5005	3·4381	3·3881
9	5·1174	4·2565	3·8626	3·6331	3·4817	3·3738	3·2927	3·2296	3·1789
10	4·9646	4·1028	3·7083	3·4780	3·3258	3·2172	3·1355	3·0717	3·0204
11	4·8443	3·9823	3·5874	3·3567	3·2039	3·0946	3·0123	2·9480	2·8962
12	4·7472	3·8853	3·4903	3·2592	3·1059	2·9961	2·9134	2·8486	2·7964
13	4·6672	3·8056	3·4105	3·1791	3·0254	2·9153	2·8321	2·7669	2·7144
14	4·6001	3·7389	3·3439	3·1122	2·9582	2·8477	2·7642	2·6987	2·6458
15	4·5431	3·6823	3·2874	3·0556	2·9013	2·7905	2·7066	2·6408	2·5876
16	4·4940	3·6337	3·2389	3·0069	2·8524	2·7413	2·6572	2·5911	2·5377
17	4·4513	3·5915	3·1968	2·9647	2·8100	2·6987	2·6143	2·5480	2·4943
18	4·4139	3·5546	3·1599	2·9277	2·7729	2·6613	2·5767	2·5102	2·4563
19	4·3808	3·5219	3·1274	2·8951	2·7401	2·6283	2·5435	2·4768	2·4227
20	4·3513	3·4928	3·0984	2·8661	2·7109	2·5990	2·5140	2·4471	2·3928
21	4·3248	3·4668	3·0725	2·8401	2·6848	2·5727	2·4876	2·4205	2·3661
22	4·3009	3·4434	3·0491	2·8167	2·6613	2·5491	2·4638	2·3965	2·3419
23	4·2793	3·4221	3·0280	2·7955	2·6400	2·5277	2·4422	2·3748	2·3201
24	4·2597	3·4028	3·0088	2·7763	2·6207	2·5082	2·4226	2·3551	2·3002
25	4·2417	3·3852	2·9912	2·7587	2·6030	2·4904	2·4047	2·3371	2·2821
26	4·2252	3·3690	2·9751	2·7426	2·5868	2·4741	2·3883	2·3205	2·2655
27	4·2100	3·3541	2·9604	2·7278	2·5719	2·4591	2·3732	2·3053	2·2501
28	4·1960	3·3404	2·9467	2·7141	2·5581	2·4453	2·3593	2·2913	2·2360
29	4·1830	3·3277	2·9340	2·7014	2·5454	2·4324	2·3463	2·2782	2·2229
30	4·1709	3·3158	2·9223	2·6896	2·5336	2·4205	2·3343	2·2662	2·2107
40	4·0848	3·2317	2·8387	2·6060	2·4495	2·3359	2·2490	2·1802	2·1240
60	4·0012	3·1504	2·7581	2·5252	2·3683	2·2540	2·1665	2·0970	2·0401
120	3·9201	3·0718	2·6802	2·4472	2·2900	2·1750	2·0867	2·0164	1·9588
∞	3·8415	2·9957	2·6049	2·3719	2·2141	2·0986	2·0096	1·9384	1·8799

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}05$.

F DISTRIBUTION: 5 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	241·88	243·91	245·95	248·01	249·05	250·09	251·14	252·20	253·25	254·32
2	19·396	19·413	19·429	19·446	19·454	19·462	19·471	19·479	19·487	19·496
3	8·7855	8·7446	8·7029	8·6602	8·6385	8·6166	8·5944	8·5720	8·5494	8·5265
4	5·9644	5·9117	5·8578	5·8025	5·7744	5·7459	5·7170	5·6878	5·6581	5·6281
5	4·7351	4·6777	4·6188	4·5581	4·5272	4·4957	4·4638	4·4314	4·3984	4·3650
6	4·0600	3·9999	3·9381	3·8742	3·8415	3·8082	3·7743	3·7398	3·7047	3·6688
7	3·6365	3·5747	3·5108	3·4445	3·4105	3·3758	3·3404	3·3043	3·2674	3·2298
8	3·3472	3·2840	3·2184	3·1503	3·1152	3·0794	3·0428	3·0053	2·9669	2·9276
9	3·1373	3·0729	3·0061	2·9365	2·9005	2·8637	2·8259	2·7872	2·7475	2·7067
10	2·9782	2·9130	2·8450	2·7740	2·7372	2·6996	2·6609	2·6211	2·5801	2·5379
11	2·8536	2·7876	2·7186	2·6464	2·6090	2·5705	2·5309	2·4901	2·4480	2·4045
12	2·7534	2·6866	2·6169	2·5436	2·5055	2·4663	2·4259	2·3842	2·3410	2·2962
13	2·6710	2·6037	2·5331	2·4589	2·4202	2·3803	2·3392	2·2966	2·2524	2·2064
14	2·6021	2·5342	2·4630	2·3879	2·3487	2·3082	2·2664	2·2230	2·1778	2·1307
15	2·5437	2·4753	2·4035	2·3275	2·2878	2·2468	2·2043	2·1601	2·1141	2·0658
16	2·4935	2·4247	2·3522	2·2756	2·2354	2·1938	2·1507	2·1058	2·0589	2·0096
17	2·4499	2·3807	2·3077	2·2304	2·1898	2·1477	2·1040	2·0584	2·0107	1·9604
18	2·4117	2·3421	2·2686	2·1906	2·1497	2·1071	2·0629	2·0166	1·9681	1·9168
19	2·3779	2·3080	2·2341	2·1555	2·1141	2·0712	2·0264	1·9796	1·9302	1·8780
20	2·3479	2·2776	2·2033	2·1242	2·0825	2·0391	1·9938	1·9464	1·8963	1·8432
21	2·3210	2·2504	2·1757	2·0960	2·0540	2·0102	1·9645	1·9165	1·8657	1·8117
22	2·2967	2·2258	2·1508	2·0707	2·0283	1·9842	1·9380	1·8895	1·8380	1·7831
23	2·2747	2·2036	2·1282	2·0476	2·0050	1·9605	1·9139	1·8649	1·8128	1·7570
24	2·2547	2·1834	2·1077	2·0267	1·9838	1·9390	1·8920	1·8424	1·7897	1·7331
25	2·2365	2·1649	2·0889	2·0075	1·9643	1·9192	1·8718	1·8217	1·7684	1·7110
26	2·2197	2·1479	2·0716	1·9898	1·9464	1·9010	1·8533	1·8027	1·7488	1·6906
27	2·2043	2·1323	2·0558	1·9736	1·9299	1·8842	1·8361	1·7851	1·7307	1·6717
28	2·1900	2·1179	2·0411	1·9586	1·9147	1·8687	1·8203	1·7689	1·7138	1·6541
29	2·1768	2·1045	2·0275	1·9446	1·9005	1·8543	1·8055	1·7537	1·6981	1·6377
30	2·1646	2·0921	2·0148	1·9317	1·8874	1·8409	1·7918	1·7396	1·6835	1·6223
40	2·0772	2·0035	1·9245	1·8389	1·7929	1·7444	1·6928	1·6373	1·5766	1·5089
60	1·9926	1·9174	1·8364	1·7480	1·7001	1·6491	1·5943	1·5343	1·4673	1·3893
120	1·9105	1·8337	1·7505	1·6587	1·6084	1·5543	1·4952	1·4290	1·3519	1·2539
∞	1·8307	1·7522	1·6664	1·5705	1·5173	1·4591	1·3940	1·3180	1·2214	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

F Distribution: 2·5 per cent Points

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9
1	647·79	799·50	864·16	899·58	921·85	937·11	948·22	956·66	963·28
2	38·506	39·000	39·165	39·248	39·298	39·331	39·355	39·373	39·387
3	17·443	16·044	15·439	15·101	14·885	14·735	14·624	14·540	14·473
4	12·218	10·649	9·9792	9·6045	9·3645	9·1973	9·0741	8·9796	8·9047
5	10·007	8·4336	7·7636	7·3879	7·1464	6·9777	6·8531	6·7572	6·6810
6	8·8131	7·2598	6·5988	6·2272	5·9876	5·8197	5·6955	5·5996	5·5234
7	8·0727	6·5415	5·8898	5·5226	5·2852	5·1186	4·9949	4·8994	4·8232
8	7·5709	6·0595	5·4160	5·0526	4·8173	4·6517	4·5286	4·4332	4·3572
9	7·2093	5·7147	5·0781	4·7181	4·4844	4·3197	4·1971	4·1020	4·0260
10	6·9367	5·4564	4·8256	4·4683	4·2361	4·0721	3·9498	3·8549	3·7790
11	6·7241	5·2559	4·6300	4·2751	4·0440	3·8807	3·7586	3·6638	3·5879
12	6·5538	5·0959	4·4742	4·1212	3·8911	3·7283	3·6065	3·5118	3·4358
13	6·4143	4·9653	4·3472	3·9959	3·7667	3·6043	3·4827	3·3880	3·3120
14	6·2979	4·8567	4·2417	3·8919	3·6634	3·5014	3·3799	3·2853	3·2093
15	6·1995	4·7650	4·1528	3·8043	3·5764	3·4147	3·2934	3·1987	3·1227
16	6·1151	4·6867	4·0768	3·7294	3·5021	3·3406	3·2194	3·1248	3·0488
17	6·0420	4·6189	4·0112	3·6648	3·4379	3·2767	3·1556	3·0610	2·9849
18	5·9781	4·5597	3·9539	3·6083	3·3820	3·2209	3·0999	3·0053	2·9291
19	5·9216	4·5075	3·9034	3·5587	3·3327	3·1718	3·0509	2·9563	2·8800
20	5·8715	4·4613	3·8587	3·5147	3·2891	3·1283	3·0074	2·9128	2·8365
21	5·8266	4·4199	3·8188	3·4754	3·2501	3·0895	2·9686	2·8740	2·7977
22	5·7863	4·3828	3·7829	3·4401	3·2151	3·0546	2·9338	2·8392	2·7628
23	5·7498	4·3492	3·7505	3·4083	3·1835	3·0232	2·9024	2·8077	2·7313
24	5·7167	4·3187	3·7211	3·3794	3·1548	2·9946	2·8738	2·7791	2·7027
25	5·6864	4·2909	3·6943	3·3530	3·1287	2·9685	2·8478	2·7531	2·6766
26	5·6586	4·2655	3·6697	3·3289	3·1048	2·9447	2·8240	2·7293	2·6528
27	5·6331	4·2421	3·6472	3·3067	3·0828	2·9228	2·8021	2·7074	2·6309
28	5·6096	4·2205	3·6264	3·2863	3·0625	2·9027	2·7820	2·6872	2·6106
29	5·5878	4·2006	3·6072	3·2674	3·0438	2·8840	2·7633	2·6686	2·5919
30	5·5675	4·1821	3·5894	3·2499	3·0265	2·8667	2·7460	2·6513	2·5746
40	5·4239	4·0510	3·4633	3·1261	2·9037	2·7444	2·6238	2·5289	2·4519
60	5·2857	3·9253	3·3425	3·0077	2·7863	2·6274	2·5068	2·4117	2·3344
120	5·1524	3·8046	3·2270	2·8943	2·6740	2·5154	2·3948	2·2994	2·2217
∞	5·0239	3·6889	3·1161	2·7858	2·5665	2·4082	2·2875	2·1918	2·1136

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}025$.

F DISTRIBUTION: 2·5 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	968·63	976·71	984·87	993·10	997·25	1001·4	1005·6	1009·8	1014·0	1018·3
2	39·398	39·415	39·431	39·448	39·456	39·465	39·473	39·481	39·490	39·498
3	14·419	14·337	14·253	14·167	14·124	14·081	14·037	13·992	13·947	13·902
4	8·8439	8·7512	8·6565	8·5599	8·5109	8·4613	8·4111	8·3604	8·3092	8·2573
5	6·6192	6·5246	6·4277	6·3285	6·2780	6·2269	6·1751	6·1225	6·0693	6·0153
6	5·4613	5·3662	5·2687	5·1684	5·1172	5·0652	5·0125	4·9589	4·9045	4·8491
7	4·7611	4·6658	4·5678	4·4667	4·4150	4·3624	4·3089	4·2544	4·1989	4·1423
8	4·2951	4·1997	4·1012	3·9995	3·9472	3·8940	3·8398	3·7844	3·7279	3·6702
9	3·9639	3·8682	3·7694	3·6669	3·6142	3·5604	3·5055	3·4493	3·3918	3·3329
10	3·7168	3·6209	3·5217	3·4186	3·3654	3·3110	3·2554	3·1984	3·1399	3·0798
11	3·5257	3·4296	3·3299	3·2261	3·1725	3·1176	3·0613	3·0035	2·9441	2·8828
12	3·3736	3·2773	3·1772	3·0728	3·0187	2·9633	2·9063	2·8478	2·7874	2·7249
13	3·2497	3·1532	3·0527	2·9477	2·8932	2·8373	2·7797	2·7204	2·6590	2·5955
14	3·1469	3·0501	2·9493	2·8437	2·7888	2·7324	2·6742	2·6142	2·5519	2·4872
15	3·0602	2·9633	2·8621	2·7559	2·7006	2·6437	2·5850	2·5242	2·4611	2·3953
16	2·9862	2·8890	2·7875	2·6808	2·6252	2·5678	2·5085	2·4471	2·3831	2·3163
17	2·9222	2·8249	2·7230	2·6158	2·5598	2·5021	2·4422	2·3801	2·3153	2·2474
18	2·8664	2·7689	2·6667	2·5590	2·5027	2·4445	2·3842	2·3214	2·2558	2·1869
19	2·8173	2·7196	2·6171	2·5089	2·4523	2·3937	2·3329	2·2695	2·2032	2·1333
20	2·7737	2·6758	2·5731	2·4645	2·4076	2·3486	2·2873	2·2234	2·1562	2·0853
21	2·7348	2·6368	2·5338	2·4247	2·3675	2·3082	2·2465	2·1819	2·1141	2·0422
22	2·6998	2·6017	2·4984	2·3890	2·3315	2·2718	2·2097	2·1446	2·0760	2·0032
23	2·6682	2·5699	2·4665	2·3567	2·2989	2·2389	2·1763	2·1107	2·0415	1·9677
24	2·6396	2·5412	2·4374	2·3273	2·2693	2·2090	2·1460	2·0799	2·0099	1·9353
25	2·6135	2·5149	2·4110	2·3005	2·2422	2·1816	2·1183	2·0517	1·9811	1·9055
26	2·5895	2·4909	2·3867	2·2759	2·2174	2·1565	2·0928	2·0257	1·9545	1·8781
27	2·5676	2·4688	2·3644	2·2533	2·1946	2·1334	2·0693	2·0018	1·9299	1·8527
28	2·5473	2·4484	2·3438	2·2324	2·1735	2·1121	2·0477	1·9796	1·9072	1·8291
29	2·5286	2·4295	2·3248	2·2131	2·1540	2·0923	2·0276	1·9591	1·8861	1·8072
30	2·5112	2·4120	2·3072	2·1952	2·1359	2·0739	2·0089	1·9400	1·8664	1·7867
40	2·3882	2·2882	2·1819	2·0677	2·0069	1·9429	1·8752	1·8028	1·7242	1·6371
60	2·2702	2·1692	2·0613	1·9445	1·8817	1·8152	1·7440	1·6668	1·5810	1·4822
120	2·1570	2·0548	1·9450	1·8249	1·7597	1·6899	1·6141	1·5299	1·4327	1·3104
∞	2·0483	1·9447	1·8326	1·7085	1·6402	1·5660	1·4835	1·3883	1·2684	1·0000

$$F=\frac{s_1^2}{s_2^2}=\frac{\nu_2 S_1}{\nu_1 S_2}.$$

F DISTRIBUTION: 1 PER CENT POINTS

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	4052·2	4999·5	5403·3	5624·6	5763·7	5859·0	5928·3	5981·6	6022·5
2	98·503	99·000	99·166	99·249	99·299	99·332	99·356	99·374	99·388
3	34·116	30·817	29·457	28·710	28·237	27·911	27·672	27·489	27·345
4	21·198	18·000	16·694	15·977	15·522	15·207	14·976	14·799	14·659
5	16·258	13·274	12·060	11·392	10·967	10·672	10·456	10·289	10·158
6	13·745	10·925	9·7795	9·1483	8·7459	8·4661	8·2600	8·1016	7·9761
7	12·246	9·5466	8·4513	7·8467	7·4604	7·1914	6·9928	6·8401	6·7188
8	11·259	8·6491	7·5910	7·0060	6·6318	6·3707	6·1776	6·0289	5·9106
9	10·561	8·0215	6·9919	6·4221	6·0569	5·8018	5·6129	5·4671	5·3511
10	10·044	7·5594	6·5523	5·9943	5·6363	5·3858	5·2001	5·0567	4·9424
11	9·6460	7·2057	6·2167	5·6683	5·3160	5·0692	4·8861	4·7445	4·6315
12	9·3302	6·9266	5·9526	5·4119	5·0643	4·8206	4·6395	4·4994	4·3875
13	9·0738	6·7010	5·7394	5·2053	4·8616	4·6204	4·4410	4·3021	4·1911
14	8·8616	6·5149	5·5639	5·0354	4·6950	4·4558	4·2779	4·1399	4·0297
15	8·6831	6·3589	5·4170	4·8932	4·5556	4·3183	4·1415	4·0045	3·8948
16	8·5310	6·2262	5·2922	4·7726	4·4374	4·2016	4·0259	3·8896	3·7804
17	8·3997	6·1121	5·1850	4·6690	4·3359	4·1015	3·9267	3·7910	3·6822
18	8·2854	6·0129	5·0919	4·5790	4·2479	4·0146	3·8406	3·7054	3·5971
19	8·1850	5·9259	5·0103	4·5003	4·1708	3·9386	3·7653	3·6305	3·5225
20	8·0960	5·8489	4·9382	4·4307	4·1027	3·8714	3·6987	3·5644	3·4567
21	8·0166	5·7804	4·8740	4·3688	4·0421	3·8117	3·6396	3·5056	3·3981
22	7·9454	5·7190	4·8166	4·3134	3·9880	3·7583	3·5867	3·4530	3·3458
23	7·8811	5·6637	4·7649	4·2635	3·9392	3·7102	3·5390	3·4057	3·2986
24	7·8229	5·6136	4·7181	4·2184	3·8951	3·6667	3·4959	3·3629	3·2560
25	7·7698	5·5680	4·6755	4·1774	3·8550	3·6272	3·4568	3·3239	3·2172
26	7·7213	5·5263	4·6366	4·1400	3·8183	3·5911	3·4210	3·2884	3·1818
27	7·6767	5·4881	4·6009	4·1056	3·7848	3·5580	3·3882	3·2558	3·1494
28	7·6356	5·4529	4·5681	4·0740	3·7539	3·5276	3·3581	3·2259	3·1195
29	7·5976	5·4205	4·5378	4·0449	3·7254	3·4995	3·3302	3·1982	3·0920
30	7·5625	5·3904	4·5097	4·0179	3·6990	3·4735	3·3045	3·1726	3·0665
40	7·3141	5·1785	4·3126	3·8283	3·5138	3·2910	3·1238	2·9930	2·8876
60	7·0771	4·9774	4·1259	3·6491	3·3389	3·1187	2·9530	2·8233	2·7185
120	6·8510	4·7865	3·9493	3·4796	3·1735	2·9559	2·7918	2·6629	2·5586
∞	6·6349	4·6052	3·7816	3·3192	3·0173	2·8020	2·6393	2·5113	2·4073

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}01$.

F DISTRIBUTION: 1 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	6055·8	6106·3	6157·3	6208·7	6234·6	6260·7	6286·8	6313·0	6339·4	6366·0
2	99·399	99·416	99·432	99·449	99·458	99·466	99·474	99·483	99·491	99·501
3	27·229	27·052	26·872	26·690	26·598	26·505	26·411	26·316	26·221	26·125
4	14·546	14·374	14·198	14·020	13·929	13·838	13·745	13·652	13·558	13·463
5	10·051	9·8883	9·7222	9·5527	9·4665	9·3793	9·2912	9·2020	9·1118	9·0204
6	7·8741	7·7183	7·5590	7·3958	7·3127	7·2285	7·1432	7·0568	6·9690	6·8801
7	6·6201	6·4691	6·3143	6·1554	6·0743	5·9921	5·9084	5·8236	5·7372	5·6495
8	5·8143	5·6668	5·5151	5·3591	5·2793	5·1981	5·1156	5·0316	4·9460	4·8588
9	5·2565	5·1114	4·9621	4·8080	4·7290	4·6486	4·5667	4·4831	4·3978	4·3105
10	4·8492	4·7059	4·5582	4·4054	4·3269	4·2469	4·1653	4·0819	3·9965	3·9090
11	4·5393	4·3974	4·2509	4·0990	4·0209	3·9411	3·8596	3·7761	3·6904	3·6025
12	4·2961	4·1553	4·0096	3·8584	3·7805	3·7008	3·6192	3·5355	3·4494	3·3608
13	4·1003	3·9603	3·8154	3·6646	3·5868	3·5070	3·4253	3·3413	3·2548	3·1654
14	3·9394	3·8001	3·6557	3·5052	3·4274	3·3476	3·2656	3·1813	3·0942	3·0040
15	3·8049	3·6662	3·5222	3·3719	3·2940	3·2141	3·1319	3·0471	2·9595	2·8684
16	3·6909	3·5527	3·4089	3·2588	3·1808	3·1007	3·0182	2·9330	2·8447	2·7528
17	3·5931	3·4552	3·3117	3·1615	3·0835	3·0032	2·9205	2·8348	2·7459	2·6530
18	3·5082	3·3706	3·2273	3·0771	2·9990	2·9185	2·8354	2·7493	2·6597	2·5660
19	3·4338	3·2965	3·1533	3·0031	2·9249	2·8442	2·7608	2·6742	2·5839	2·4893
20	3·3682	3·2311	3·0880	2·9377	2·8594	2·7785	2·6947	2·6077	2·5168	2·4212
21	3·3098	3·1729	3·0299	2·8796	2·8011	2·7200	2·6359	2·5484	2·4568	2·3603
22	3·2576	3·1209	2·9780	2·8274	2·7488	2·6675	2·5831	2·4951	2·4029	2·3055
23	3·2106	3·0740	2·9311	2·7805	2·7017	2·6202	2·5355	2·4471	2·3542	2·2559
24	3·1681	3·0316	2·8887	2·7380	2·6591	2·5773	2·4923	2·4035	2·3099	2·2107
25	3·1294	2·9931	2·8502	2·6993	2·6203	2·5383	2·4530	2·3637	2·2695	2·1694
26	3·0941	2·9579	2·8150	2·6640	2·5848	2·5026	2·4170	2·3273	2·2325	2·1315
27	3·0618	2·9256	2·7827	2·6316	2·5522	2·4699	2·3840	2·2938	2·1984	2·0965
28	3·0320	2·8959	2·7530	2·6017	2·5223	2·4397	2·3535	2·2629	2·1670	2·0642
29	3·0045	2·8685	2·7256	2·5742	2·4946	2·4118	2·3253	2·2344	2·1378	2·0342
30	2·9791	2·8431	2·7002	2·5487	2·4689	2·3860	2·2992	2·2079	2·1107	2·0062
40	2·8005	2·6648	2·5216	2·3689	2·2880	2·2034	2·1142	2·0194	1·9172	1·8047
60	2·6318	2·4961	2·3523	2·1978	2·1154	2·0285	1·9360	1·8363	1·7263	1·6006
120	2·4721	2·3363	2·1915	2·0346	1·9500	1·8600	1·7628	1·6557	1·5330	1·3805
∞	2·3209	2·1848	2·0385	1·8783	1·7908	1·6964	1·5923	1·4730	1·3246	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

F Distribution: 0·5 per cent Points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	16211	20000	21615	22500	23056	23437	23715	23925	24091
2	198·50	199·00	199·17	199·25	199·30	199·33	199·36	199·37	199·39
3	55·552	49·799	47·467	46·195	45·392	44·838	44·434	44·126	43·882
4	31·333	26·284	24·259	23·155	22·456	21·975	21·622	21·352	21·139
5	22·785	18·314	16·530	15·556	14·940	14·513	14·200	13·961	13·772
6	18·635	14·544	12·917	12·028	11·464	11·073	10·786	10·566	10·391
7	16·236	12·404	10·882	10·050	9·5221	9·1554	8·8854	8·6781	8·5138
8	14·688	11·042	9·5965	8·8051	8·3018	7·9520	7·6942	7·4960	7·3386
9	13·614	10·107	8·7171	7·9559	7·4711	7·1338	6·8849	6·6933	6·5411
10	12·826	9·4270	8·0807	7·3428	6·8723	6·5446	6·3025	6·1159	5·9676
11	12·226	8·9122	7·6004	6·8809	6·4217	6·1015	5·8648	5·6821	5·5368
12	11·754	8·5096	7·2258	6·5211	6·0711	5·7570	5·5245	5·3451	5·2021
13	11·374	8·1865	6·9257	6·2335	5·7910	5·4819	5·2529	5·0761	4·9351
14	11·060	7·9217	6·6803	5·9984	5·5623	5·2574	5·0313	4·8566	4·7173
15	10·798	7·7008	6·4760	5·8029	5·3721	5·0708	4·8473	4·6743	4·5364
16	10·575	7·5138	6·3034	5·6378	5·2117	4·9134	4·6920	4·5207	4·3838
17	10·384	7·3536	6·1556	5·4967	5·0746	4·7789	4·5594	4·3893	4·2535
18	10·218	7·2148	6·0277	5·3746	4·9560	4·6627	4·4448	4·2759	4·1410
19	10·073	7·0935	5·9161	5·2681	4·8526	4·5614	4·3448	4·1770	4·0428
20	9·9439	6·9865	5·8177	5·1743	4·7616	4·4721	4·2569	4·0900	3·9564
21	9·8295	6·8914	5·7304	5·0911	4·6808	4·3931	4·1789	4·0128	3·8799
22	9·7271	6·8064	5·6524	5·0168	4·6088	4·3225	4·1094	3·9440	3·8116
23	9·6348	6·7300	5·5823	4·9500	4·5441	4·2591	4·0469	3·8822	3·7502
24	9·5513	6·6610	5·5190	4·8898	4·4857	4·2019	3·9905	3·8264	3·6949
25	9·4753	6·5982	5·4615	4·8351	4·4327	4·1500	3·9394	3·7758	3·6447
26	9·4059	6·5409	5·4091	4·7852	4·3844	4·1027	3·8928	3·7297	3·5989
27	9·3423	6·4885	5·3611	4·7396	4·3402	4·0594	3·8501	3·6875	3·5571
28	9·2838	6·4403	5·3170	4·6977	4·2996	4·0197	3·8110	3·6487	3·5186
29	9·2297	6·3958	5·2764	4·6591	4·2622	3·9830	3·7749	3·6130	3·4832
30	9·1797	6·3547	5·2388	4·6233	4·2276	3·9492	3·7416	3·5801	3·4505
40	8·8278	6·0664	4·9759	4·3738	3·9860	3·7129	3·5088	3·3498	3·2220
60	8·4946	5·7950	4·7290	4·1399	3·7600	3·4918	3·2911	3·1344	3·0083
120	8·1790	5·5393	4·4973	3·9207	3·5482	3·2849	3·0874	2·9330	2·8083
∞	7·8794	5·2983	4·2794	3·7151	3·3499	3·0913	2·8968	2·7444	2·6210

This table gives the values of F for which $I_F(\nu_1, \nu_2) = 0{\cdot}005$.

F DISTRIBUTION: 0·5 PER CENT POINTS

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465
2	199·40	199·42	199·43	199·45	199·46	199·47	199·47	199·48	199·49	199·51
3	43·686	43·387	43·085	42·778	42·622	42·466	42·308	42·149	41·989	41·829
4	20·967	20·705	20·438	20·167	20·030	19·892	19·752	19·611	19·468	19·325
5	13·618	13·384	13·146	12·903	12·780	12·656	12·530	12·402	12·274	12·144
6	10·250	10·034	9·8140	9·5888	9·4741	9·3583	9·2408	9·1219	9·0015	8·8793
7	8·3803	8·1764	7·9678	7·7540	7·6450	7·5345	7·4225	7·3088	7·1933	7·0760
8	7·2107	7·0149	6·8143	6·6082	6·5029	6·3961	6·2875	6·1772	6·0649	5·9505
9	6·4171	6·2274	6·0325	5·8318	5·7292	5·6248	5·5186	5·4104	5·3001	5·1875
10	5·8467	5·6613	5·4707	5·2740	5·1732	5·0705	4·9659	4·8592	4·7501	4·6385
11	5·4182	5·2363	5·0489	4·8552	4·7557	4·6543	4·5508	4·4450	4·3367	4·2256
12	5·0855	4·9063	4·7214	4·5299	4·4315	4·3309	4·2282	4·1229	4·0149	3·9039
13	4·8199	4·6429	4·4600	4·2703	4·1726	4·0727	3·9704	3·8655	3·7577	3·6465
14	4·6034	4·4281	4·2468	4·0585	3·9614	3·8619	3·7600	3·6553	3·5473	3·4359
15	4·4236	4·2498	4·0698	3·8826	3·7859	3·6867	3·5850	3·4803	3·3722	3·2602
16	4·2719	4·0994	3·9205	3·7342	3·6378	3·5388	3·4372	3·3324	3·2240	3·1115
17	4·1423	3·9709	3·7929	3·6073	3·5112	3·4124	3·3107	3·2058	3·0971	2·9839
18	4·0305	3·8599	3·6827	3·4977	3·4017	3·3030	3·2014	3·0962	2·9871	2·8732
19	3·9329	3·7631	3·5866	3·4020	3·3062	3·2075	3·1058	3·0004	2·8908	2·7762
20	3·8470	3·6779	3·5020	3·3178	3·2220	3·1234	3·0215	2·9159	2·8058	2·6904
21	3·7709	3·6024	3·4270	3·2431	3·1474	3·0488	2·9467	2·8408	2·7302	2·6140
22	3·7030	3·5350	3·3600	3·1764	3·0807	2·9821	2·8799	2·7736	2·6625	2·5455
23	3·6420	3·4745	3·2999	3·1165	3·0208	2·9221	2·8198	2·7132	2·6016	2·4837
24	3·5870	3·4199	3·2456	3·0624	2·9667	2·8679	2·7654	2·6585	2·5463	2·4276
25	3·5370	3·3704	3·1963	3·0133	2·9176	2·8187	2·7160	2·6088	2·4960	2·3765
26	3·4916	3·3252	3·1515	2·9685	2·8728	2·7738	2·6709	2·5633	2·4501	2·3297
27	3·4499	3·2839	3·1104	2·9275	2·8318	2·7327	2·6296	2·5217	2·4078	2·2867
28	3·4117	3·2460	3·0727	2·8899	2·7941	2·6949	2·5916	2·4834	2·3689	2·2469
29	3·3765	3·2111	3·0379	2·8551	2·7594	2·6601	2·5565	2·4479	2·3330	2·2102
30	3·3440	3·1787	3·0057	2·8230	2·7272	2·6278	2·5241	2·4151	2·2997	2·1760
40	3·1167	2·9531	2·7811	2·5984	2·5020	2·4015	2·2958	2·1838	2·0635	1·9318
60	2·9042	2·7419	2·5705	2·3872	2·2898	2·1874	2·0789	1·9622	1·8341	1·6885
120	2·7052	2·5439	2·3727	2·1881	2·0890	1·9839	1·8709	1·7469	1·6055	1·4311
∞	2·5188	2·3583	2·1868	1·9998	1·8983	1·7891	1·6691	1·5325	1·3637	1·0000

$$F = \frac{s_1^2}{s_2^2} = \frac{\nu_2 S_1}{\nu_1 S_2}.$$

The accompanying tables give seven upper percentage points for F, that is they give the roots of the equation

$$I_F(\nu_1, \nu_2) = \int_F^\infty f(F)\, dF \qquad (10)$$

for $100 I_F(\nu_1, \nu_2) = 50, 25, 10, 5, 2{\cdot}5, 1$ and $0{\cdot}5$,

and for $\nu_1 = 1(1)10, 12, 15, 20, 24, 30, 40, 60, 120$ and ∞

$\nu_2 = 1(1)30, 40, 60, 120$ and ∞.

To obtain the lower percentage points, that is the roots of

$$I'_F(\nu_1, \nu_2) = \int_0^F f(F)\, dF = I_{1/F}(\nu_2, \nu_1),$$

it is only necessary to interchange the values of ν_1 and ν_2 in entering the tables and to take for F the reciprocal of the value so obtained. For example if $\nu_1 = 12$, $\nu_2 = 40$, the upper 0·5 % point is seen to be

$$F_{0{\cdot}005}(12, 40) = 2{\cdot}9531.$$

The lower 0·5 % point is

$$F_{0{\cdot}995}(12, 40) = 1/F_{0{\cdot}005}(40, 12) = 1/4{\cdot}2282 = 0{\cdot}2365.$$

The marginal row of each table under the heading $\nu_2 = \infty$ provides the corresponding upper percentage point of the distribution of χ^2/ν with $\nu = \nu_1$ degrees of freedom. The marginal column of each table for $\nu_1 = \infty$ gives the upper percentage point of ν/χ^2, or its reciprocal gives the lower percentage point of χ^2/ν with $\nu = \nu_2$ degrees of freedom. The calculations involved in forming these columns and rows were the basis of the tables of percentage points of χ^2 recently published in this journal (Catherine M. Thompson, 1941 *b*).

The percentage points for the 'Student' ratio t having $\nu = \nu_2$ degrees of freedom may be obtained from the column of the tables headed $\nu_1 = 1$, since in this case $t = \sqrt{F}$. This relation was used to form the short table of percentage points of the t-distribution published in the last issue of this journal (Merrington, 1942).

The entries in the main body of the tables were computed by Mrs Maxine Merrington from Miss Thompson's (1941 *a*) values of the percentage points of x, using the transformation

$$F = \frac{p(1-x)}{qx}.$$

The results are given to five significant figures.

In issuing the tables in present form the question of the number of figures to be retained needed some consideration. For the ordinary purpose of significance tests there is no doubt that the two-decimal place accuracy usually given for F (e.g. Snedecor (1934) and Fisher & Yates (1938)) is ample, and in a book of tables issued primarily for the working statistician this consideration would be the deciding factor. Experience, however, has shown that the table-maker can never be sure of the purposes for which the table-user may need his work, either now or at some future date. For example, Simaika (1942) has recently discussed methods of interpolating additional percentage points between tabulated pivotal values. For this purpose considerable accuracy is required for the latter values. It was therefore decided to publish these tables to the full accuracy available both in the direct (x) and inverted (F) forms of the beta distribution.

The methods of interpolation discussed by Hartley (1941) and Comrie & Hartley (1941) for the case of the x-tables will in general be applicable also for the F-tables.

REFERENCES

Comrie, L. J. & Hartley, H. O. (1941). *Biometrika*, **32**, 183–6.

Fisher, R. A. & Yates, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research.* Table V. Oliver and Boyd.

Hartley, H. O. (1941). *Biometrika*, **32**, 161–7.

Merrington, Maxine (1942). *Biometrika*, **32**, 300.

Pearson, K. (1934). *Tables of the Incomplete Beta-function.* London: Biometrika.

Simaika, J. B. (1942). *Biometrika*, **32**, 263–76.

Snedecor, G. W. (1934). *Calculation and Interpretation of Analysis of Variance and Covariance.* Ames, Iowa: Collegiate Press Inc.

Thompson, Catherine M. (1941 *a*). *Biometrika*, **32**, 168–81.

Thompson, Catherine M. (1941 *b*). *Biometrika*, **32**, 188–9.

TABLES OF THE PROBABILITY INTEGRAL OF THE STUDENTIZED RANGE

By E. S. PEARSON AND H. O. HARTLEY

1. Introduction

Denote by $x_1, \ldots, x_n$ a random sample of n observations arranged in ascending order of magnitude and drawn from a normal population with standard deviation σ. The range, or spread, in the sample is then $x_n - x_1$ and may be expressed in units of the standard deviation by the ratio

$$w = \frac{x_n - x_1}{\sigma}. \tag{1}$$

Tables of the probability integral of w were given in the last issue of *Biometrika* (Hartley, 1942; Pearson, 1942), together with tables of its percentage points.* The uses of the range cover a wide field. As an example of particular importance we should mention the setting up of quality control charts in industry, where the simplicity of range is of great advantage.

The probability distribution of the difference $x_n - x_1$ depends on the standard deviation σ, in the sampled population. In many cases this will be unknown, but can be estimated from a second independent sample $X_1, \ldots, X_{\nu+1}$, drawn from the same population. If we denote by

$$s^2 = \frac{1}{\nu} \sum_{1}^{\nu+1} (X - \bar{X})^2,$$

the estimate of σ^2 derived therefrom, it is easy to see that the probability law of the ratio

$$q = \frac{x_n - x_1}{s} \tag{2}$$

does not depend on σ.

This idea of 'studentizing' the range seems to have occurred first to W. S. Gosset himself (see letter of 29 January 1932 quoted by E. S. Pearson, 1938, p. 245). Following this suggestion, D. Newman (1939) calculated a number of percentage points for q. Newman's table was obtained by quadrature from E. S. Pearson's approximate probability law of w. Since exact tables of this latter integral are now available, it appeared appropriate to revise and amplify the 'studentized' distribution law resulting from it. Moreover, certain new results (to be published separately) make it possible to simplify both the calculation of 'studentized' probability laws as well as their tabulations. It suffices here to *state* these results as applied to the range.

2. Description of the tables

The probability law of q depends both on the size n of the first sample (from which the range $x_n - x_1$ is determined) and on the degrees of freedom ν of the standard deviation estimated from the second sample. The probability integral may be denoted by ${}_\nu P_n(Q)$ and represents the chance that the ratio q does not exceed the limit Q. As $\nu \to \infty$ and $s^2 \to \sigma^2$, ${}_\nu P_n(Q)$ will tend to the probability integral $P_n(W)$ of the ratio $w = (x_n - x_1)/\sigma$, taken at $W = Q$.

* See also Simon (1941, pp. 204–7). There are slight, though practically unimportant, inaccuracies in his Table C 2.

It can now be proved that to an accuracy sufficient for practical purposes and for $\nu \geq 10$, ${}_\nu P_n(Q)$ may be represented as a quadratic in $1/\nu$, viz.

$$ {}_\nu P_n(Q) = P_n(Q) + \frac{1}{\nu} a_n(Q) + \frac{1}{\nu^2} b_n(Q). \tag{3}$$

The coefficients $P_n(Q)$, $a_n(Q)$ and $b_n(Q)$ are given in Table 1, and the following example illustrates their use.

Example. Consider the range in samples of 12 and an independent estimate of the standard deviation based on 15 degrees of freedom. Find the chance that in random sampling the former exceeds 4 times the latter.

Entering Table 1 for $n = 12$ and $Q = 4$, we find

$$P_{12}(4) = 0\cdot8321, \quad a_{12}(4) = -1\cdot71, \quad b_{12}(4) = 6\cdot2,$$

whence $${}_{15}P_{12}(4) = 0\cdot8321 - \frac{1\cdot71}{15} + \frac{6\cdot2}{(15)^2} = 0\cdot8321 - 0\cdot114 + 0\cdot028 = 0\cdot746,$$

so that the required chance is equal to

$$1 - {}_{15}P_{12}(4) = 0\cdot254.$$

If the desired value of Q is not a tabular value, ${}_\nu P_n(Q)$ is found by interpolation Q-wise.

For certain applications it is necessary to know the limits Q corresponding to standard probability levels. Table 2 gives these percentage points for four levels, viz. the upper and lower 5 and 1 % points.*

The lower percentage points are given to two places of decimals, for degrees of freedom ν, ranging from 10 to ∞ and, for sample sizes n, between 2 and 20. The upper percentage points are given for the same values of ν and n and to the same accuracy, except that for $10 \leq \nu < 20$ only one place of decimals is given. This drop in accuracy is necessary as formula (3) is less accurate for small values of ν and large values of Q. One decimal accuracy in the percentage points is, however, considered ample having regard to the inaccuracies, which may be considerable for large Q, introduced through possible deviations from normality in the parental distribution. When using these tables it is advisable, therefore, to judge percentage points whose values exceed 6 with discretion.

A comparison of the upper percentage points given in Table 2 with those given in Newman's table (1939, p. 25) shows that while the latter may be in error by as much as 1 or 2 units in the first decimal, it provides a useful working guide to the significance levels of q.

3. Applications

We confine ourselves to a selected number of applications without claiming to cover the whole field. The first example is a modification of one used by 'Student' (1927, pp. 161–2).

Example 1. *Control of accuracy in chemical routine analysis.* The problem which 'Student' considered was the common one with which the industrial chemist is faced of making day after day a certain number of similar routine analyses of some solution or substance that must be regularly checked for conformity to standard. The characteristic measured, for example, the acidity of a solution, is estimated from the mean of a few (say n) observations, and a routine check on the consistency of these determinations is required to ensure that

* These levels are most commonly used in tests of significance, but other levels may be required. For instance, the 2·5 % point and the 0·1 % point are frequently used as 'inner' and 'outer' limits on quality control charts. Such additional percentage points can, of course, be calculated from Table 1 by inverse interpolation.

accuracy is maintained. Discordant observations will be repeated and, if necessary, rejected, and 'Student' pointed out that it would obviously be of advantage to work on a regular system; for this purpose he proposed the use of the range of the n determinations. The situation he considered was one in which the standard deviation of the within-day error of analysis had been found from experience to remain constant and could be assigned a known value, σ.

It is clear, however, that situations will occur in which the standard error of analysis appropriate on a given day can only be estimated from the determinations of a few previous days, perhaps because a new chemist has been put on to the work or because the method of analysis has changed. Instead, therefore, of basing the regular check system on the values of $w = (x_n - x_1)/\sigma$, as 'Student' did, we shall suppose it to be based on $q = (x_n - x_1)/s$; we then have the following modification of his example.

On each day four determinations are made in the first instance; if these show too great a scatter the data have to be improved by additional tests. Suppose that it is considered advisable to base s, the estimate of the appropriate σ, on the tests of five previous days only, i.e. on 20 observations, then ν will equal 15. In the units dealt with, let us suppose we obtain $s = 0{\cdot}675$. Then the 5 % points of Table 2 (for 15 degrees of freedom) provide gauge values for the actual range $x_n - x_1$, as follows:

Size of sample n	Gauge value $Q_n s$ for the actual range $R_n = x_n - x_1$
4	$0{\cdot}675 \times 4{\cdot}1 = 2{\cdot}8 = Q_4 s$
5	$\times 4{\cdot}4 = 3{\cdot}0 = Q_5 s$
6	$\times 4{\cdot}6 = 3{\cdot}1 = Q_6 s$
7	$\times 4{\cdot}8 = 3{\cdot}2 = Q_7 s$

'Student's' procedure is now as follows. If the actual range of a day's sample of four test results (R_4) is greater than $Q_4 s$, repetition of the test should be made giving rise to a sample of 5 with range R_5. If R_5 is smaller than $Q_5 s$ the mean of the five results should be accepted as the day's mean. If, on the other hand, R_5 exceeds $Q_5 s$ the most outlying observation should be rejected and if the resulting sample of 4 has a range R_4 smaller than $Q_4 s$ the mean of these four tests should be accepted; but if not, a further repetition should be made and the whole sample of six tests examined afresh, and so on until a sample of at least four tests with R_n smaller than $Q_n s$ is obtained.

Suppose, for example, that for a particular day the four test results are 22·8, 23·5, 26·0 and 26·6. We find that the range R_4 is 3·8 and exceeds $Q_4 s$. We therefore repeat the test, obtaining (say) 23·9 as the fifth result. For our sample of 5, $R_5 = 3{\cdot}8$, which exceeds $Q_5 s$, and we reject the result 22·8. This leaves us a sample of 4 with $R_4 = 3{\cdot}1$, which is still in excess of $Q_4 s$. We therefore repeat again, getting (say) 23·5. Now we have a sample of 6 with a range of $R_6 = 3{\cdot}8$, which exceeds $Q_6 s$. We therefore reject 22·8 again, to reach a sample of 5 with $R_5 = 3{\cdot}1$, and rejecting 26·6 a sample of 4 with $R_4 = 2{\cdot}5$. This is smaller than $Q_4 s$ and therefore we accept the mean (24·2) of the remaining test results (23·5, 26·0, 23·9, 23·5) as the day's mean.

In this particular example it happens that there is little alteration in the procedure when compared with 'Student's' procedure, which was based on the percentage points of w. In general it may be said that if there is sufficient information to show that σ remains constant from day to day, it is preferable to use a long-term estimate based on many degrees of freedom, so that the percentage points of q assume the limiting values of those for w, i.e. the values in the bottom row of Table 2 with $\nu = \infty$. To take account only of the last few

days' results in estimating σ would lead to unnecessary latitude in the gauge we are using. If, however, there are reasons for supposing that σ may change, then we must use the short-term estimate and the resulting greater uncertainty requires the wider q-limits.

In conclusion we should quote 'Student': 'It should always be remembered that such rules are to be regarded as aids to and not as substitutes for common sense!'

Example 2. *Control charts for range.* The preceding example was concerned with controlling the accuracy of a testing technique. In the usual quality control problem, charts for variability are usually concerned with real fluctuations in the quality of the manufactured article. Here the use of range in the place of standard deviation or mean deviation has the advantage of simplicity. For a full description of applications the reader is referred to Dudding & Jennett (1942), Pearson (1935) or Simon (1941).

The variability in the quality of manufactured products will often be stable and may then be adequately represented by a fixed known standard deviation σ (which in some cases is fixed by specification). With some processes, however, variability in quality is influenced by short-term factors. Although such fluctuation may be well within the permissible tolerance limits, the standard deviation will have to be estimated from a limited number of immediately preceding observations, i.e. the control limits must be based on s rather than σ and latitude must be allowed for consequent greater uncertainty. Here the percentage points for q rather than w are appropriate.

As an example consider the manufacture, on a quantity basis, of a component of an electro-mechanical instrument which has to be machined to fairly close tolerances. Suppose that the component is produced by a battery of eight machines and that during each shift samples of five components are measured (to precision) for each of the eight machines. A control chart for range is to be used to control the variability in general and the uniformity of the average performance of the eight machines in particular.

Slow secular changes in the standard deviation of the measurements are observed and a fresh estimate of the standard deviation is, therefore, calculated weekly from the first batch of eight samples of five measurements. Suppose that the standard deviation, s, within samples (based on 32 degrees of freedom) is 0·0015 inch; using now the upper and lower percentage points, we obtain by interpolation the values corresponding to $n = 5$ and $\nu = 32$ and find for the 1 % control limits:

$$\text{Upper limit} = 5{\cdot}03 \times 0{\cdot}0015 = 0{\cdot}0075 \text{ inch,}$$
$$\text{Lower limit} = 0{\cdot}66 \times 0{\cdot}0015 = 0{\cdot}0010 \text{ inch.}$$

It is now desired to control the uniformity in average performance of the eight machines. To this end we may calculate for each machine the average of the five measurements* in each sample tested and plot, on a second chart, the range of the eight 'machine averages'. The control limits will be at

$$\text{Upper limit} = 5{\cdot}51 \times \frac{0{\cdot}0015}{\sqrt{5}} = 0{\cdot}0037 \text{ inch,}$$

$$\text{Lower limit} = 1{\cdot}17 \times \frac{0{\cdot}0015}{\sqrt{5}} = 0{\cdot}0008 \text{ inch.}$$

Example 3. *A special problem of 'spread' in machine part assembly.* In the above examples the range was used as a simple short-cut measure of the variability of test results. It was chosen because of its simplicity rather than on theoretical grounds. There are, however,

* These averages have to be computed in any case, as they are required for the control chart for mean.

problems where the requirements of the applications (usually dictated by certain tolerance limits) demand a control of range. Instances such as the spread of salvos from a battery of guns or the spread in a stick of bombs cannot be dealt with here. We shall discuss, however, a rather special example taken from the assembly of electromagnetic machine parts and stated in a simplified form.

Electromagnetic relays are manufactured to a specified setting-up time. This is the time elapsing between the primary impulse in the coil and the complete contact in the secondary circuit of the relay. For each individual relay the actual setting-up time may differ from specification but will stay practically constant in time. Sets of (say) 15 relays are now assembled in a machine. To prevent 'arcing' a cam must break the secondary circuit before the first relay has set up until after the last relay has set up. The length of the break interval is a fixed characteristic of the cam and samples of 15 relays whose range in setting-up times exceeds this interval cannot, therefore, be fitted into the machine. Thus, on testing, the slowest or the fastest relay will have to be replaced and it is necessary to keep the frequency of such replacements below a reasonably low percentage.

Table 1 gives this frequency. Suppose, for example, in the first place that we have enough information to regard the standard deviation of the setting-up times as known and equal to 1/25th of a second; further, that the cam has a break-interval of 1/5th of a second. Then $Q = 0{\cdot}2/0{\cdot}04 = 5$, $n = 15$ and we find from the table that $P_{15}(5) = 0{\cdot}9688$. Thus we see that the expected frequency for replacements is

$$100(1 - P_{15}(5)) = 3{\cdot}12\ \%.$$

In general the formulae:

$$Q = \frac{\text{Cam-break-interval}}{\text{Standard deviation of setting-up times}}$$

$$100(1 - P_n(Q)) = \text{Percentage frequency of necessary replacements,}$$

$$n = \text{Number of relays assembled in the machine,}$$

relate the accuracy of the relays to the cam-break-interval. The table may, therefore, be used as a guide when deciding on tolerance limits in the manufacture of relays or in designing a cam to 'make it fit' the relays.

If the standard deviation of the setting-up times is not known exactly, but has to be estimated from a limited sample, the integral $P_n(Q)$ has to be replaced by ${}_\nu P_n(Q)$, which is given by formula (3).

Further applications of the tables to the Analysis of Variance of field experiments are given in Newman's (1939) paper.

4. Calculation of tables

Both tables are based on a five-decimal manuscript table of the probability integral of range $P_n(W)$, four decimals of which were published in the last issue of *Biometrika* (Pearson, 1942). Of the coefficients given in Table 1, $P_n(Q)$ was copied from the published table, putting $Q = W$. The formulae for the coefficients a_n and b_n are

$$a_n(Q) = \frac{1}{4}\left\{Q^2\frac{d^2P_n}{dW^2} - Q\frac{dP_n}{dW}\right\},$$

$$b_n(Q) = \frac{1}{16}\left\{\frac{Q^4}{2}\frac{d^4P_n}{dW^4} - \frac{Q^3}{3}\frac{d^3P_n}{dW^3} - \frac{Q^2}{2}\frac{d^2P_n}{dW^2} + Q\frac{dP_n}{dW}\right\}.$$

The derivatives in these formulae were calculated from the differences of $P_n(W)$ (at interval 0·25) by standard formulae and are taken at argument $W = Q$. Checks consisted in differencing Q-wise and n-wise, but owing to rounding-off errors in the differences the last figure of $b_n(Q)$ cannot be guaranteed near the bottom of Table 1. A special marginal check was obtained from 'Student's' table of the probability integral of t (1925), using the relation $\sqrt{2} \times t = q$ for $n = 2$.

Table 2 was produced from Table 1 by inverse interpolation using the relation (3).

We wish to acknowledge the careful work of Mrs M. Merrington, who carried out most of the calculations.

REFERENCES

DUDDING, B. P. & JENNETT, W. J. (1942). *The Application of Statistical Methods to Quality Control.* British Standards Institution, No. 600 (revise).

HARTLEY, H. O. (1942). *Biometrika*, **32**, 334.

NEWMAN, D. (1939). *Biometrika*, **31**, 20.

PEARSON, E. S. (1935). *The Application of Statistical Methods to Industrial Standardization and Quality Control.* British Standards Institution, No. 600.

—— (1938). *Biometrika*, **30**, 210.

—— (1942). *Biometrika*, **32**, 301.

SIMON, L. E. (1941). *An Engineer's Manual of Statistical Methods.* New York: John Wiley and Sons, Inc.

'STUDENT' (1927). *Biometrika*, **19**, 151.

—— (1925). *Metron*, **5**, 105.

Table 1. *For calculating the probability integral of* $q = \dfrac{x_n - x_1}{s}$

$Q \backslash n$	P_n	a_n	b_n	P_n	a_n	b_n	P_n	a_n	b_n
		3			4			5	
0·00	0·0000			0·0000			0·0000		
0·25	·0171			·0020			·0002		
0·50	·0666			·0152	+0·01		·0033	+0·01	
0·75	·1436	−0·02		·0483	+0·02		·0157	+0·02	
1·00	0·2407	−0·06		0·1057	+0·03		0·0450	+0·05	−0·1
1·25	·3495	−0·12		·1868	0·00		·0970	+0·07	−0·1
1·50	·4614	−0·21	+0·1	·2865	−0·06		·1733	+0·06	−0·1
1·75	·5690	−0·31	+0·2	·3970	−0·17	+0·1	·2706	−0·02	−0·1
2·00	0·6665	−0·40	+0·3	0·5096	−0·32	+0·4	0·3816	−0·16	+0·2
2·25	·7505	−0·48	+0·4	·6163	−0·47	+0·6	·4968	−0·36	+0·6
2·50	·8195	−0·53	+0·4	·7110	−0·61	+0·9	·6075	−0·57	+1·1
2·75	·8737	−0·54	+0·3	·7905	−0·69	+1·0	·7063	−0·74	+1·5
3·00	0·9145	−0·51	+0·1	0·8537	−0·72	+0·8	0·7891	−0·85	+1·6
3·25	·9439	−0·46	−0·2	·9016	−0·70	+0·5	·8546	−0·88	+1·3
3·50	·9644	−0·39	−0·4	·9361	−0·69	−0·1	·9037	−0·84	+0·7
3·75	·9782	−0·31	−0·7	·9600	−0·52	−0·6	·9386	−0·74	−0·2
4·00	0·9870	−0·24	−0·9	0·9758	−0·42	−1·1	0·9623	−0·61	−0·9
4·25	·9925	−0·18	−1·0	·9859	−0·31	−1·4	·9777	−0·47	−1·6
4·50	·9958	−0·12	−1·0	·9920	−0·22	−1·6	·9873	−0·34	−1·9
4·75	·9977	−0·08	−0·9	·9956	−0·15	−1·5	·9930	−0·24	−1·9
5·00	0·9988	−0·05	−0·7	0·9977	−0·10	−1·3	0·9963	−0·16	−1·8
5·25	·9994	−0·03	−0·5	·9988	−0·06	−1·0	·9981	−0·10	−1·6
5·50	·9997	−0·02	−0·4	·9994	−0·04	−0·8	·9990	−0·06	−1·2
5·75	·9999	−0·01	−0·3	·9997	−0·02	−0·6	·9995	−0·03	−0·9
6·00	0·9999	0·00	−0·2	0·9999	−0·01	−0·4	0·9998	−0·02	−0·6

$Q \backslash n$	P_n	a_n	b_n	P_n	a_n	b_n	P_n	a_n	b_n
		6			7			8	
0·50	0·0007			0·0002			0·0000		
0·75	·0050	+0·02		·0016	+0·01		·0005		
1·00	0·0188	+0·05		0·0078	+0·04		0·0032	+0·02	
1·25	·0495	+0·09	−0·1	·0250	+0·08	−0·1	·0124	+0·07	
1·50	·1031	+0·13	−0·2	·0606	+0·15	−0·2	·0353	+0·14	−0·2
1·75	·1815	+0·11	−0·3	·1204	+0·19	−0·4	·0792	+0·22	−0·5
2·00	0·2816	+0·01	−0·2	0·2056	+0·15	−0·5	0·1489	+0·25	−0·7
2·25	·3955	−0·19	+0·2	·3118	−0·01	−0·3	·2440	+0·15	−0·7
2·50	·5132	−0·44	+1·0	·4300	−0·27	+0·6	·3579	−0·09	0·0
2·75	·6252	−0·70	+1·7	·5494	−0·59	+1·6	·4800	−0·44	+1·2
3·00	0·7239	−0·90	+2·3	0·6601	−0·88	+2·7	0·5991	−0·81	+2·8
3·25	·8053	−1·02	+2·2	·7553	−1·09	+3·1	·7055	−1·10	+3·7
3·50	·8685	−1·02	+1·6	·8316	−1·16	+2·7	·7938	−1·26	+3·9
3·75	·9148	−0·94	+0·7	·8891	−1·12	+1·6	·8622	−1·28	+2·8
4·00	0·9469	−0·80	−0·6	0·9300	−0·99	+0·1	0·9120	−1·16	+1·0
4·25	·9682	−0·63	−1·5	·9576	−0·81	−1·3	·9461	−0·98	−0·7
4·50	·9817	−0·47	−2·1	·9754	−0·61	−2·4	·9684	−0·76	−2·2
4·75	·9898	−0·33	−2·4	·9862	−0·44	−3·0	·9822	−0·55	−3·4
5·00	0·9946	−0·22	−2·5	0·9926	−0·30	−3·0	0·9903	−0·38	−3·4
5·25	·9972	−0·14	−2·1	·9961	−0·19	−2·8	·9949	−0·25	−3·4
5·50	·9986	−0·09	−1·7	·9981	−0·12	−2·2	·9974	−0·15	−2·9
5·75	·9993	−0·05	−1·3	·9991	−0·07	−1·7	·9988	−0·09	−2·2
6·00	0·9997	−0·03	−0·9	0·9996	−0·04	−1·2	0·9994	−0·05	−1·5

$$_{\nu}P_n(Q) = P_n(Q) + \frac{1}{\nu}a_n(Q) + \frac{1}{\nu^2}b_n(Q)$$

Table 1 (cont.). *For calculating the probability integral of* $q = \frac{x_n - x_1}{s}$

Q \ n	P_n (9)	a_n (9)	b_n (9)	P_n (10)	a_n (10)	b_n (10)	P_n (11)	a_n (11)	b_n (11)
0·75	0·0002			0·0000			0·0000		
1·00	0·0013	+0·01		0·0005	+0·01		0·0002		
1·25	·0062	+0·05		·0030	+0·03		·0015	+0·02	+0·1
1·50	·0204	+0·12	−0·1	·0117	+0·10		·0067	+0·07	+0·1
1·75	·0517	+0·23	−0·4	·0336	+0·21	−0·3	·0217	+0·18	−0·1
2·00	0·1072	+0·30	−0·8	0·0768	+0·33	−0·8	0·0548	+0·33	−0·7
2·25	·1899	+0·28	−1·1	·1470	+0·37	−1·4	·1134	+0·43	−1·5
2·50	·2964	+0·08	−0·6	·2443	+0·24	−1·3	·2007	+0·37	−1·8
2·75	·4175	−0·26	+0·6	·3617	−0·08	−0·2	·3124	+0·10	−1·0
3·00	0·5415	−0·68	+2·6	0·4878	−0·53	+2·1	0·4382	−0·36	+1·4
3·25	·6569	−1·06	+4·2	·6099	−0·99	+4·3	·5649	−0·88	+4·1
3·50	·7558	−1·32	+4·9	·7180	−1·34	+5·6	·6807	−1·32	+6·3
3·75	·8345	−1·40	+4·0	·8062	−1·50	+5·4	·7776	−1·57	+6·6
4·00	0·8929	−1·33	+2·2	0·8731	−1·47	+3·4	0·8528	−1·60	+4·8
4·25	·9338	−1·14	+0·1	·9208	−1·31	+1·0	·9072	−1·46	+2·2
4·50	·9608	−0·91	−2·1	·9527	−1·06	−1·8	·9441	−1·20	−1·2
4·75	·9777	−0·67	−3·6	·9729	−0·80	−3·5	·9678	−0·92	−3·3
5·00	0·9878	−0·47	−4·0	0·9851	−0·56	−4·6	0·9822	−0·65	−5·0
5·25	·9936	−0·31	−3·9	·9922	−0·37	−4·4	·9906	−0·44	−5·0
5·50	·9968	−0·19	−3·5	·9960	−0·23	−4·1	·9952	−0·28	−4·6
5·75	·9984	−0·11	−2·7	·9980	−0·14	−3·3	·9976	−0·16	−3·8
6·00	0·9993	−0·07	−1·8	0·9991	−0·08	−2·3	0·9989	−0·09	−2·8
6·25	·9997	−0·03	−1·3	·9996	−0·04	−1·7	·9995	−0·05	−1·9
6·50	·9998	−0·02	−0·8	·9998	−0·02	−0·9	·9998	−0·03	−1·1

Q \ n	P_n (12)	a_n (12)	b_n (12)	P_n (13)	a_n (13)	b_n (13)	P_n (14)	a_n (14)	b_n (14)
1·00	0·0001			0·0000			0·0000		
1·25	·0007	+0·01	+0·1	·0004	+0·01		·0002		
1·50	·0038	+0·05	+0·1	·0022	+0·04	+0·1	·0012	+0·03	+0·1
1·75	·0140	+0·15	0·0	·0090	+0·12	+0·1	·0058	+0·10	+0·2
2·00	0·0389	+0·31	−0·5	0·0276	+0·28	−0·3	0·0195	+0·24	−0·1
2·25	·0872	+0·46	−1·5	·0669	+0·46	−1·4	·0511	+0·45	−1·1
2·50	·1644	+0·47	−2·2	·1342	+0·55	−2·4	·1094	+0·59	−2·5
2·75	·2690	+0·27	−1·8	·2311	+0·41	−2·5	·1981	+0·54	−3·3
3·00	0·3927	−0·18	+0·5	0·3512	0·0	−0·5	0·3134	+0·17	−1·6
3·25	·5222	−0·74	+3·7	·4817	−0·59	+3·0	·4437	−0·42	+2·1
3·50	·6442	−1·26	+6·7	·6087	−1·18	+6·8	·5744	−1·08	+6·7
3·75	·7491	−1·61	+7·6	·7206	−1·61	+8·5	·6925	−1·60	+9·2
4·00	0·8321	−1·71	+6·2	0·8111	−1·80	+7·7	0·7899	−1·86	+9·1
4·25	·8931	−1·60	+3·4	·8787	−1·73	+4·7	·8639	−1·85	+6·2
4·50	·9352	−1·35	−0·5	·9258	−1·49	+0·5	·9162	−1·63	+1·6
4·75	·9624	−1·05	−3·1	·9567	−1·18	−2·7	·9508	−1·30	−2·2
5·00	0·9791	−0·75	−5·2	0·9759	−0·85	−5·3	0·9724	−0·96	−5·3
5·25	·9889	−0·51	−5·6	·9871	−0·58	−6·0	·9852	−0·66	−6·3
5·50	·9943	−0·32	−5·0	·9934	−0·37	−5·6	·9924	−0·42	−6·3
5·75	·9972	−0·19	−4·3	·9967	−0·22	−4·8	·9962	−0·26	−5·4
6·00	0·9987	−0·11	−3·1	0·9984	−0·13	−3·7	0·9982	−0·15	−4·2
6·25	·9994	−0·06	−2·2	·9993	−0·07	−2·6	·9992	−0·08	−2·9
6·50	·9997	−0·03	−1·5	·9997	−0·03	−1·6	·9996	−0·04	−1·8

$${}_\nu P_n(Q) = P_n(Q) + \frac{1}{\nu} a_n(Q) + \frac{1}{\nu^2} b_n(Q)$$

Table 1 (cont.). *For calculating the probability integral of* $q = \dfrac{x_n - x_1}{s}$

Q \ n	P_n (15)	a_n (15)	b_n (15)	P_n (16)	a_n (16)	b_n (16)	P_n (17)	a_n (17)	b_n (17)
1·00	0·0000			0·0000			0·0000		
1·25	·0001			·0000			·0000		
1·50	·0007	+0·02	+ 0·1	·0004	+0·01	+ 0·1	·0002	+0·01	+ 0·1
1·75	·0037	+0·07	+ 0·2	·0024	+0·06	+ 0·2	·0015	+0·04	+ 0·2
2·00	0·0137	+0·21	0·0	0·0097	+0·18	+ 0·2	0·0068	+0·14	+ 0·3
2·25	·0390	+0·42	− 0·9	·0297	+0·39	− 0·6	·0226	+0·35	− 0·4
2·50	·0890	+0·62	− 2·5	·0722	+0·62	− 2·3	·0586	+0·62	− 2·0
2·75	·1696	+0·64	− 3·8	·1448	+0·72	− 4·1	·1236	+0·77	− 4·3
3·00	0·2792	+0·34	− 2·6	0·2484	+0·49	− 3·6	0·2207	+0·62	− 4·5
3·25	·4081	−0·25	+ 1·0	·3748	−0·07	− 0·2	·3438	+0·10	− 1·4
3·50	·5413	−0·96	+ 6·2	·5096	−0·82	+ 5·5	·4792	−0·67	+ 4·7
3·75	·6648	−1·56	+ 9·8	·6376	−1·50	+10·1	·6110	−1·42	+10·1
4·00	0·7686	−1·91	+10·4	0·7474	−1·94	+11·6	0·7263	−1·95	+12·8
4·25	·8488	−1·96	+ 7·7	·8336	−2·05	+ 9·2	·8182	−2·13	+10·7
4·50	·9062	−1·76	+ 2·8	·8960	−1·88	+ 4·1	·8856	−2·00	+ 5·5
4·75	·9446	−1·43	− 1·6	·9383	−1·55	− 0·8	·9317	−1·68	+ 0·2
5·00	0·9688	−1·06	− 5·3	0·9650	−1·17	− 5·2	0·9611	−1·28	− 5·0
5·25	·9832	−0·73	− 6·8	·9811	−0·82	− 7·1	·9789	−0·90	− 7·3
5·50	·9913	−0·48	− 6·8	·9902	−0·53	− 7·5	·9890	−0·59	− 8·1
5·75	·9957	−0·29	− 5·8	·9951	−0·33	− 6·4	·9945	−0·37	− 7·1
6·00	0·9979	−0·17	− 4·7	0·9976	−0·19	− 5·3	0·9974	−0·21	− 5·8
6·25	·9990	−0·09	− 3·4	·9989	−0·11	− 3·7	·9988	−0·12	− 4·1
6·50	·9996	−0·05	− 2·1	·9995	−0·05	− 2·3	·9995	−0·06	− 2·6

Q \ n	P_n (18)	a_n (18)	b_n (18)	P_n (19)	a_n (19)	b_n (19)	P_n (20)	a_n (20)	b_n (20)
1·00	0·0000			0·0000			0·0000		
1·25	·0000			·0000			·0000		
1·50	·0001		+ 0·1	·0001		+ 0·1	·0000		
1·75	·0010	+0·03	+ 0·2	·0006	+0·02	+ 0·2	·0004	+0·02	+ 0·2
2·00	0·0048	+0·12	+ 0·4	0·0033	+0·10	+ 0·4	0·0023	+0·08	+ 0·4
2·25	·0172	+0·32	− 0·1	·0130	+0·28	+ 0·2	·0099	+0·24	+ 0·4
2·50	·0474	+0·59	− 1·7	·0383	+0·56	− 1·4	·0309	+0·53	− 1·0
2·75	·1053	+0·81	− 4·4	·0896	+0·83	− 4·3	·0761	+0·83	− 4·0
3·00	0·1959	+0·73	− 5·2	0·1736	+0·83	− 5·9	0·1538	+0·91	− 6·4
3·25	·3151	+0·27	− 2·7	·2884	+0·42	− 4·0	·2638	+0·57	− 5·2
3·50	·4502	−0·51	+ 3·6	·4226	−0·35	+ 2·4	·3964	−0·18	+ 1·1
3·75	·5850	−1·33	+ 9·9	·5598	−1·22	+ 9·6	·5352	−1·10	+ 8·9
4·00	0·7053	−1·94	+13·7	0·6845	−1·92	+14·5	0·6640	−1·88	+15·2
4·25	·8027	−2·19	+12·2	·7871	−2·25	+13·7	·7715	−2·29	+15·1
4·50	·8750	−2·11	+ 7·1	·8643	−2·21	+ 8·7	·8534	−2·30	+10·2
4·75	·9249	−1·80	+ 1·2	·9180	−1·91	+ 2·3	·9110	−2·03	+ 3·6
5·00	0·9571	−1·38	− 4·7	0·9529	−1·49	− 4·2	0·9486	−1·59	− 3·7
5·25	·9766	−0·98	− 7·4	·9742	−1·07	− 7·4	·9718	−1·16	− 7·3
5·50	·9878	−0·65	− 8·5	·9865	−0·71	− 8·7	·9852	−0·77	− 8·9
5·75	·9939	−0·40	− 7·9	·9932	−0·44	− 8·4	·9925	−0·48	− 8·8
6·00	0·9971	−0·24	− 6·3	0·9967	−0·26	− 6·9	0·9964	−0·29	− 7·4
6·25	·9986	−0·13	− 4·6	·9985	−0·14	− 5·1	·9983	−0·16	− 5·5
6·50	·9994	−0·07	− 3·0	·9993	−0·08	− 3·3	·9992	−0·08	− 3·6

$$_{\nu}P_n(Q) = P_n(Q) + \frac{1}{\nu}\, a_n(Q) + \frac{1}{\nu^2}\, b_n(Q)$$

Table 2. *Lower percentage points of the studentized range* $q = \frac{x_n - x_1}{s}$.

5 % points

ν \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
10	0·09	0·43	0·75	1·01	1·20	1·37	1·52	1·63	1·74	1·83	1·91	1·98	2·05	2·12	2·17	2·22	2·26	2·30	2·34
11	·09	·43	·75	·01	·21	·38	·52	·64	·75	·84	·92	2·00	·07	·13	·18	·24	·28	·33	·37
12	·09	·43	·75	·01	·21	·38	·53	·65	·76	·85	·93	·01	·08	·14	·20	·26	·30	·34	·38
13	·09	·43	·75	·01	·22	·39	·53	·65	·76	·86	·94	·02	·09	·15	·21	·27	·31	·36	·40
14	·09	·43	·75	·01	·22	·39	·54	·66	·77	·86	·95	·03	·10	·16	·22	·28	·32	·37	·41
15	0·09	0·43	0·75	1·01	1·22	1·39	1·54	1·66	1·77	1·87	1·95	2·03	2·11	2·17	2·23	2·29	2·34	2·38	2·43
16	·09	·43	·75	·01	·22	·39	·54	·67	·78	·87	·96	·04	·11	·18	·24	·30	·34	·39	·44
17	·09	·43	·75	·01	·22	·40	·55	·67	·78	·88	·97	·05	·12	·19	·25	·30	·35	·40	·45
18	·09	·43	·75	·02	·22	·40	·55	·67	·79	·88	·97	·05	·12	·19	·25	·31	·36	·41	·45
19	·09	·43	·75	·02	·23	·40	·55	·68	·79	·89	·98	·05	·13	·20	·26	·32	·37	·42	·46
20	0·09	0·43	0·75	1·02	1·23	1·40	1·55	1·68	1·79	1·89	1·98	2·06	2·13	2·20	2·27	2·32	2·37	2·42	2·47
24	·09	·43	·75	·02	·23	·41	·56	·69	·80	·90	·99	·08	·15	·22	·28	·34	·39	·45	·49
30	·09	·43	·76	·02	·24	·41	·57	·70	·81	·92	2·01	·09	·17	·24	·30	·36	·41	·47	·52
40	·09	·43	·76	·02	·24	·42	·57	·71	·82	·93	·02	·10	·18	·26	·32	·38	·43	·49	·54
60	0·09	0·43	0·76	1·02	1·24	1·43	1·58	1·72	1·83	1·94	2·04	2·12	2·20	2·28	2·34	2·40	2·46	2·52	2·57
120	·09	·43	·76	·03	·25	·43	·59	·73	·85	·96	·06	·14	·22	·30	·36	·43	·49	·54	·60
∞	0·09	0·43	0·76	1·03	1·25	1·44	1·60	1·74	1·86	1·97	2·07	2·16	2·24	2·32	2·39	2·45	2·52	2·57	2·62

1 % points

ν \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
10	0·02	0·18	0·42	0·64	0·81	0·96	1·11	1·23	1·34	1·41	1·50	1·57	1·62	1·70	1·74	1·81	1·84	1·88	1·92
11	·02	·18	·42	·64	·82	·97	·12	·24	·35	·43	·52	·58	·64	·71	·76	·82	·86	·91	·94
12	·02	·18	·42	·64	·82	·98	·12	·24	·35	·44	·53	·60	·65	·73	·77	·84	·88	·92	·96
13	·02	·18	·42	·64	·83	·98	·13	·25	·36	·45	·54	·61	·66	·74	·79	·85	·89	·94	·98
14	·02	·18	·42	·65	·83	·99	·13	·25	·37	·46	·55	·62	·68	·76	·80	·87	·91	·95	2·00
15	0·02	0·18	0·42	0·65	0·83	0·99	1·14	1·26	1·37	1·46	1·55	1·63	1·69	1·76	1·81	1·88	1·92	1·97	2·01
16	·02	·18	·42	·65	·83	·99	·14	·26	·37	·47	·56	·63	·70	·77	·82	·89	·93	·98	·02
17	·02	·18	·42	·65	·84	1·00	·14	·27	·38	·48	·57	·64	·70	·78	·83	·90	·94	·99	·04
18	·02	·18	·42	·65	·84	·00	·15	·27	·38	·48	·57	·65	·71	·79	·84	·91	·95	2·00	·05
19	·02	·18	·43	·65	·84	·00	·15	·28	·39	·48	·58	·65	·72	·80	·85	·91	·96	·01	·06
20	0·02	0·18	0·43	0·65	0·84	1·01	1·15	1·28	1·39	1·49	1·58	1·66	1·72	1·80	1·85	1·92	1·97	2·01	2·06
24	·02	·18	·43	·65	·85	·01	·16	·29	·40	·50	·60	·67	·74	·82	·88	·94	·99	·05	·09
30	·02	·18	·43	·66	·85	·02	·17	·30	·41	·52	·61	·69	·76	·84	·90	·97	2·02	·07	·12
40	·02	·18	·43	·66	·85	·02	·18	·31	·43	·53	·63	·71	·79	·86	·92	·99	·04	·10	·15
60	0·02	0·18	0·43	0·66	0·86	1·03	1·19	1·32	1·44	1·55	1·64	1·73	1·81	1·88	1·95	2·02	2·07	2·13	2·18
120	·02	·18	·43	·66	·86	·04	·20	·33	·45	·56	·66	·75	·83	·91	·98	·04	·10	·16	·21
∞	0·02	0·19	0·43	0·66	0·87	1·05	1·20	1·34	1·47	1·58	1·68	1·77	1·86	1·93	2·01	2·08	2·14	2·20	2·25

Table 2 (cont.). *Upper percentage points of the studentized range* $q = \frac{x_n - x_1}{s}$.

5 % points

ν \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
10	3·15	3·9	4·4	4·8	5·0	5·3	5·4	5·6	5·7	5·8	5·9	6·0	6·1	6·1	6·2	6·3	6·4	6·4	6·5
11	·11	·8	·3	·6	4·9	·1	·3	·5	·6	·7	8	5·9	·0	·0	·1	·2	·3	·3	·4
12	·08	·8	·2	·6	·8	·1	·2	·4	·5	·6	·7	·8	5·9	·0	·0	·1	·2	·2	·3
13	·06	·7	·2	·5	·8	·0	·1	·3	·4	·5	·6	·7	·8	5·9	·0	·0	·1	·1	·2
14	·03	·7	·1	·4	·7	4·9	·1	·2	·3	·5	·6	·7	·7	·8	5·9	·0	·0	·1	·1
15	3·01	3·7	4·1	4·4	4·6	4·8	5·0	5·2	5·3	5·4	5·5	5·6	5·7	5·7	5·8	5·9	6·0	6·0	6·1
16	·00	·7	·1	·4	·6	·8	·0	·1	·2	·3	·4	·5	·6	·7	·8	·8	5·9	·0	·0
17	2·98	·6	·0	·3	·5	·8	4·9	·1	·2	·3	·4	·5	·6	·6	·7	·8	·8	5·9	·0
18	·97	·6	·0	·3	·5	·7	·9	·0	·1	·2	·3	·4	·5	·6	·7	·7	·8	·9	5·9
19	·96	·6	·0	·3	·5	·7	·8	·0	·1	·2	·3	·4	·5	·5	·6	·7	·7	·8	·8
20	2·95	3·58	3·97	4·25	4·46	4·65	4·80	4·94	5·06	5·16	5·25	5·34	5·42	5·50	5·57	5·63	5·69	5·75	5·80
24	·92	·53	·90	·18	·38	·55	·70	·83	4·95	·05	·14	·22	·29	·36	·43	·49	·55	·60	·65
30	·89	·49	·85	·11	·30	·47	·61	·73	·84	4·93	·02	·10	·17	·23	·30	·36	·41	·46	·50
40	·86	·44	·79	·04	·23	·39	·52	·64	·74	·83	4·91	4·98	·05	·11	·17	·23	·28	·33	·37
60	2·83	3·40	3·74	3·98	4·16	4·32	4·44	4·55	4·65	4·73	4·81	4·88	4·94	5·00	5·06	5·11	5·16	5·20	5·24
120	·80	·36	·69	·92	·10	·24	·36	·47	·56	·64	·71	·78	·84	4·90	4·95	·00	·04	·09	·13
∞	2·77	3·32	3·63	3·86	4·03	4·17	4·29	4·39	4·47	4·55	4·62	4·68	4·74	4·80	4·84	4·89	4·93	4·97	5·01

1 % points

ν \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
10	4·48	5·1	5·6	5·9	6·1	6·3	6·4	6·5	6·6	6·7	6·8	6·8	6·9	6·9	7·0	7·0	7·0	7·1	7·1
11	·39	·0	·5	·8	·0	·2	·3	·4	·5	·6	·7	·7	·8	·9	6·9	6·9	·0	·0	·1
12	·32	·0	·4	·7	5·9	·1	·2	·4	·5	·6	·6	·7	·7	·8	·9	·9	6·9	·0	·0
13	·26	4·9	·3	·6	·8	·0	·1	·3	·4	·5	·6	·6	·7	·7	·8	·8	·8	6·9	·0
14	·21	·9	·3	·6	·8	5·9	·1	·2	·4	·5	·5	·6	·6	·7	·7	·8	·8	·9	6·9
15	4·17	4·8	5·2	5·5	5·7	5·9	6·0	6·2	6·3	6·4	6·5	6·5	6·6	6·6	6·7	6·7	6·8	6·8	6·9
16	·13	·8	·2	·4	·7	·8	·0	·1	·2	·4	·4	·5	·5	·6	·6	·7	·7	·8	·8
17	·10	·7	·1	·4	·6	·8	5·9	·1	·2	·3	·3	·4	·5	·5	·6	·6	·7	·7	·8
18	·07	·7	·1	·4	·6	·7	·9	·0	·1	·2	·3	·4	·4	·5	·5	·6	·6	·7	·7
19	·05	·7	·1	·3	·5	·7	·8	·0	·1	·2	·2	·3	·4	·4	·5	·5	·6	·6	·7
20	4·02	4·65	5·02	5·30	5·51	5·67	5·80	5·93	6·04	6·12	6·20	6·27	6·34	6·40	6·45	6·51	6·56	6·61	6·66
24	3·96	·55	4·91	·17	·38	·54	·68	·80	5·90	5·98	·07	·14	·20	·26	·32	·38	·43	·48	·52
30	·89	·46	·81	·05	·24	·40	·54	·65	·76	·83	5·92	5·99	·05	·12	·19	·24	·28	·33	·37
40	·82	·37	4·69	4·94	·11	·26	·40	·51	·60	·68	·76	·83	5·90	5·96	·02	·07	·12	·16	·21
60	3·76	4·29	4·59	4·81	4·99	5·13	5·26	5·36	5·44	5·53	5·60	5·67	5·73	5·78	5·84	5·89	5·94	5·98	6·02
120	·70	·20	·50	·71	·87	·01	·12	·22	·30	·37	·44	·51	·56	·61	·66	·71	·75	·79	5·83
∞	3·64	4·12	4·40	4·60	4·76	4·88	4·99	5·08	5·16	5·23	5·29	5·35	5·40	5·45	5·49	5·53	5·57	5·61	5·64

MISCELLANEA

Minimum range for quasi-normal distributions

By R. C. GEARY

For one variate a *quasi-normal** distribution $f(x)\,dx$ is defined as follows:

(i) $f(x)$ is continuous for $-\infty \leqslant x \leqslant +\infty$;

(ii) the distribution has a single mode;

(iii) for all values of x less than the mode $f'(x)$ is non-negative, and for all values of x greater than the mode $f'(x)$ is negative.

It will first be shown that corresponding to a given probability α, the range of y, from x to $x+y$, is shortest when

$$f(x) = f(x+y). \tag{1}$$

The property is almost obvious from geometrical considerations; it may be well, however, to give an algebraic proof:

$$1-\alpha = \int_x^{x+y} f(x)\,dx. \tag{2}$$

Differentiating, $$0 = f(x+y)\,(dx+dy) - f(x)\,dx.$$

Hence $$\frac{dy}{dx} = \frac{f(x)}{f(x+y)} - 1, \tag{3}$$

which assumes a limiting value for $dy/dx = 0$ or

$$f(x) = f(x+y).$$

Also, from (3), $$\frac{d^2y}{dx^2} = \{f(x+y)f'(x) - f'(x+y)f(x)\}/\{f(x+y)\}^2, \tag{4}$$

when $dy/dx = 0$. From (ii) above it is clear that the right side of (4) is positive. Hence $dy/dx = 0$ defines a minimum value of y in terms of x. From (1) and (2) the values of x and y are theoretically determinable.

This property can readily be generalized. A quasi-normal multiple frequency distribution

$$f(x_1, x_2, \ldots, x_k)\,dx_1\,dx_2 \ldots dx_k,$$

is defined as follows:

(i) $f(x_1, x_2, \ldots, x_k)$ is continuous in each of the variates from $-\infty$ to $+\infty$;

(ii) the distribution has a single mode;

(iii) each distribution in one dimension linearly derivable from $f(x_1, x_2, \ldots, x_k)$ is quasi-normal in the sense explained above for one variate, i.e. the distribution on the straight line $L_i(x_1, x_2, \ldots, x_k) = 0$, $i = 1, 2, \ldots, k-1$, where the L_i are of one dimension in the x_j, is quasi-normal. It will be shown that of all the 'surfaces' Σ which satisfy

$$1-\alpha = \int_k \cdots \int_\Sigma f(x_1, x_2, \ldots, x_k)\,dx_1\,dx_2 \ldots dx_k, \tag{5}$$

the surface of smallest 'volume' is given by

$$f(x_1, x_2, \ldots, x_k) = C, \quad \text{a constant.} \tag{6}$$

In (5) integration is extended to the inside of Σ which, like (6), is assumed to be closed.

Suppose, in fact, that at any two points $E'(x_1', x_2', \ldots, x_k')$, $E''(x_1'' x_2'', \ldots, x_k'')$ on Σ, $f(E') \neq f(E'')$, while Σ has its minimum volume. By an orthogonal change of variables x into Z the new Z_k axis can be made parallel to the line joining E' and E''. In the first stage of the integration of the right side of (5) one element will consist of

$$dZ_1\,dZ_2 \ldots dZ_{k-1} \int_{Z_k'}^{Z_k''} F(Z_1^0, Z_2^0, \ldots, Z_{k-1}^0, Z_k)\,dZ_k, \tag{7}$$

where F is the transform of f and the points E' and E'' have become respectively $(Z_1^0, Z_2^0, \ldots, Z_{k-1}^0, Z_k')$ and $(Z_1^0, Z_2^0, \ldots, Z_{k-1}^0, Z_k'')$. From (iii) and since $F(E') \neq F(E'')$ it is clear that without changing

* It is a pity that the term *normal* cannot be applied to this system (which includes the great majority of distributions met with in practice), the term *Gaussian* being reserved exclusively for the distribution generally known as 'normal' by writers in English.

$Z_1^0, Z_2^0, \ldots, Z_{k-1}^0$ or $dZ_1, dZ_2, \ldots, dZ_{k-1}$ two values of Z_k, say Y_k' and Y_k'' different from Z_k' or Z_k'', can be found, so that the value of (7) is unchanged but for which

$$|Y_k' - Y_k''| < |Z_k' - Z_k''|. \tag{8}$$

The surface Σ can then be continuously distorted to pass through the points $(Z_1^0, Z_2^0, \ldots, Z_{k-1}^0, Y_k')$ and $(Z_1^0, Z_2^0, \ldots, Z_{k-1}^0, Y_k'')$ with the obvious result that its volume will be diminished which is contrary to hypothesis. Hence at each point on Σ

$$f(x_1, x_2, \ldots, x_k) = C, \tag{9}$$

where the constant C is derivable from (5).

This property gives particular significance to the concept which Karl Pearson (1900) used in the derivation of the χ^2-test. Pearson's idea of integrating the $(k-1)$-dimensional integral

$$C \int \ldots \int_\Sigma e^{-\frac{1}{2}\chi^2} d\Omega = 1 - \alpha, \tag{10}$$

where
$$\chi^2 = \sum_{i=1}^{k} \frac{(x_i - \bar{x}_i)^2}{\bar{x}_i},$$

with
$$\sum_{i=1}^{k} (x_i - \bar{x}_i) = 0,$$

within the surface $\chi^2 =$ constant, was simply a brilliant mathematical device for reducing the probability integral at (10) from $k-1$ to one dimension. There is an infinity of surfaces Σ which have the property indicated at (10). It will now be observed that of all the possible surfaces, that which has the smallest value of

$$\int \ldots \int_\Sigma d\Omega$$

in the plane $\Sigma x_i = \Sigma \bar{x}_i$ is $\chi^2 =$ constant, the constant value being determined by (10).

In connexion with the estimation from random samples of parameters entering into the parent distribution, there is the analogous problem of determining the shortest fiducial or confidence interval corresponding to a probability α. In the case of one unknown parameter, Neyman (1937) has shown that statements can be made in the form

$$F_1(x_1, x_2, \ldots, x_k) \leqslant \theta \leqslant F_2(x_1, x_2, \ldots, x_k),$$

where the observed sample is $x_1, x_2, \ldots, x_k$ and where the functions F_1 and F_2 are determinable from the parent distribution function $f(x, \theta)$, in the sense that if the experiment be repeated many times the statement will be true in approximately 100 $(1-\alpha)$ % of cases and false in about 100α % of cases, α having some value like 0·05, 0·01, etc., determined in advance. If a single sufficient statistic y (a function of $x_1, x_2, \ldots, x_k$) is available in respect of θ, the statement assumes a very simple form. The limits of range of θ, namely θ_1 and θ_2, are given by the equations

$$\int_{-\infty}^{X} \phi(y, \theta_1)\, dy = \xi, \quad \int_{X}^{\infty} \phi(y, \theta_2)\, dy = \eta, \quad \xi + \eta = \alpha, \tag{11}$$

when $\phi(y, \theta)$, the distribution function of y, has the property that, corresponding to each pair of positive quantities ξ and η consistent with $\xi + \eta = \alpha$ given in advance, θ_1 and θ_2 are monotonic functions of X, the sample value of y. Theoretically θ_1 and θ_2 are determinable from (11) in terms of ξ and it would appear that the solution required is the value of ξ which minimizes $|\theta_1 - \theta_2|$. Actually the problem is more complicated because $|\theta_1 - \theta_2|$ may not be the most suitable definition of range. For example, $|\log \theta_1 - \log \theta_2|$ might be taken and, in general, the value of ξ which minimizes $|\theta_1 - \theta_2|$ will be different from that which minimizes $|\log \theta_1 - \log \theta_2|$. Neyman (1937) attempted to avoid the difficulty by adopting a probabilistic concept but he shows that this particular concept is generally inapplicable. In the following applications the metrical standard is used.

I. *Parameter of position*

Suppose that the parameter is one of position only. Equations (11) then become

$$\int_{-\infty}^{x} f(x - \theta_1)\, dx = \xi, \quad \int_{x}^{\infty} f(x - \theta_2)\, dx = \eta, \quad \xi + \eta = \alpha, \tag{12}$$

where x is the sample value of the sufficient statistic.

Set $z = \theta_1 - \theta_2$. Then, since $d\xi + d\eta = 0$,

$$\frac{dz}{d\theta_1} = 1 - \frac{f(x-\theta_1)}{f(x-\theta_2)}, \tag{13}$$

and

$$\frac{d^2z}{d\theta_1^2} = \{f(x-\theta_2)f'_x(x-\theta_1) - f(x-\theta_1)f'_x(x-\theta_2)\}/\{f(x-\theta_2)\}^2. \tag{14}$$

From (13) $dz/d\theta_1 = 0$ gives

$$f(x-\theta_1) = f(x-\theta_2), \tag{15}$$

which from (14) shows a minimum solution when $f'_x(x-\theta_1) > 0$ and $f'_x(x-\theta_2) < 0$. In the Gaussian case

$$2x = \theta_1 + \theta_2. \tag{16}$$

Equations (12) and (15) (or, in the Gaussian case, (16)) will give the values of θ_1 and θ_2 corresponding to the probability α.

II. *Estimation of m from t*

For a normal sample of n the mean of the universe m will be estimated from the Gosset-Fisher function t. The mean $\bar{x}$ and the variance s^2 having been given from the random sample of n, the requisite equations are

$$C\int_{-\infty}^{x}\left(1+\frac{t^2}{n-1}\right)^{-\frac{1}{2}n} dt = \xi, \quad C\int_{y}^{\infty}\left(1+\frac{t^2}{n-1}\right)^{-\frac{1}{2}n} dt = \eta, \quad \xi+\eta = \alpha, \tag{17}$$

with

$$x = (\bar{x}-m_1)\sqrt{n}/s, \quad y = (\bar{x}-m_2)\sqrt{n}/s.$$

Set $z = m_1 - m_2 = s(y-x)/\sqrt{n}$.

From $d\xi + d\eta = 0$ and setting $dz/dx = 0$, on reduction,

$$z = -2xs/\sqrt{n},$$

which is equivalent to

$$2\bar{x} = m_1 + m_2. \tag{18}$$

From (17) and (18), m_1 and m_2 are determinable.

III. *Estimation of the variance*

The estimation of the variance from a Gaussian sample of n illustrates the difficulty referred to above about finding a suitable standard for determining the shortest interval. In the Gaussian case the universal mean determines the position of the curve and the variance its scale, so that intuitively one feels that $|\theta_1 - \theta_2|$ is a suitable measure of range for the first and θ_1/θ_2 or $|\log\theta_1 - \log\theta_2|$ is the best measure for the second. The equations are then as follows:

$$C\int_0^v e^{-(n-1)w/2\theta_1}\left(\frac{w}{\theta_1}\right)^{\frac{1}{2}(n-3)} d\left(\frac{w}{\theta_1}\right) = \xi, \quad C\int_v^{\infty} e^{-(n-1)w/2\theta_2}\left(\frac{w}{\theta_2}\right)^{\frac{1}{2}(n-3)} d\left(\frac{w}{\theta_2}\right) = \eta, \quad \text{with} \quad \xi+\eta = \alpha, \tag{19}$$

where v is the variance calculated from the random sample of n. The constant C depends on n only.

Set

$$\frac{v}{\theta_1} = x, \quad \frac{v}{\theta_2} = y \quad \text{and} \quad z = \frac{y}{x} = \frac{\theta_1}{\theta_2}. \tag{20}$$

The conditon $d\xi + d\eta = 0$ gives

$$\frac{dz}{dx} = \left[\exp\left\{-\frac{n-1}{2}x(1-z)\right\} - z^{\frac{1}{2}(n-1)}\right] \Big/ xz^{\frac{1}{2}(n-3)},$$

so that, for $dz/dx = 0$,

$$x(z-1) = \log z. \tag{21}$$

To find an approximate solution of this equation set

$$z = 1 + t/\sqrt{n}, \quad x = 1 - u/\sqrt{n}.$$

Then t can be expanded as follows (to $n^{-\frac{3}{2}}$)

$$t = 2u + \tfrac{8}{3}u^2n^{-\frac{1}{2}} + \tfrac{28}{9}u^3n^{-1} + \tfrac{464}{135}u^4n^{-\frac{3}{2}} + \dots, \tag{22}$$

or

$$z = 1 + 2(1-x) + \tfrac{8}{3}(1-x)^2 + \tfrac{28}{9}(1-x)^3 + \tfrac{464}{135}(1-x)^4 + \dots, \tag{23}$$

correct to $n^{-\frac{3}{2}}$. Comparison with

$$z' = x^{-2} = (1-\overline{1-x})^{-2} = 1 + 2(1-x) + 3(1-x)^2 + 4(1-x)^3 + 5(1-x)^4 + \dots, \tag{24}$$

shows that, to $n^{-\frac{1}{2}}$,

$$z = x^{-2},$$

or

$$v^2 = \theta_1\theta_2 \text{ approximately.} \tag{25}$$

From (19) and (23) (or approximately by (25)), θ_1 and θ_2 giving minimum θ_1/θ_2 can be determined. The analogy will be evident between (18), where the sample mean appears as the arithmetic mean of the two limits of estimate, and (25) where the sample variance is the geometric mean of the limits of estimate. The former relation is, however, absolute while the latter is only approximate.

REFERENCES

Neyman, J. (1937). *Philos. Trans.* A, **236**, 333–80.

Pearson, K. (1900). *Phil. Mag.* **50**, 157.

BOOKS RECEIVED

The Fundamental Principles of Mathematical Statistics, with special reference to the requirements of Actuaries and Vital Statisticians, and an outline of a course in graduation. By HUGH H. WOLFENDEN. Published for the Actuarial Society of America, New York, by the Macmillan Company of Canada Ltd., Toronto. 1942. Price $5.00.

Industrial Statistics. Statistical technique applied to problems in industrial research and quality control. By H. A. FREEMAN. New York: John Wiley and Sons, Inc. London: Chapman and Hall, Ltd. 1942. Price $2.50.

The Adolescent Criminal. A medico-sociological study of 4000 male adolescents. By W. NORWOOD EAST in collaboration with PERCY STOCKS and H. T. P. YOUNG, with a Foreword by Sir ALEXANDER MAXWELL. London: J. & A. Churchill, Ltd. 1942. Price 45*s*.

Breathing Capacity and Grip Strength of Preschool Children. By ELEANOR METHENY. Published by the University of Iowa Press, Iowa City, Iowa. 1940.

Infant and Maternal Mortality, in relation to size of family and rapidity of breeding. A study in human responsibility. By C. M. BURNS, with a Foreword by the Rt Hon. Lord EUSTACE PERCY. From the Department of Physiology, King's College, University of Durham. Price 5*s*.

Year Book of Labour Statistics. Sixth year of issue, 1941. Montreal: International Labour Office. 1942. Price 8*s*.

Vol. XXXIII. Part II August, 1944

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

ASSISTED BY

EGON S. PEARSON

Reprinted by offset-litho

ISSUED BY THE BIOMETRIC LABORATORY
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

ON AUTOREGRESSIVE TIME SERIES

BY M. G. KENDALL

1. The product-moment correlation coefficients obtained by correlating the members of a time-series among themselves provide a useful method of investigating the behaviour of the series, especially in regard to oscillatory movements. Consider a series of values $x_1 \dots x_n$, measured about their mean, trend-free and defined at equal intervals of time $t = 1, \dots, n$. The kth serial correlation r_k is defined as the correlation between members of the series k intervals apart,

$$r_k = \frac{\sum_{j=1}^{n-k} x_j x_{j+k}}{\left\{\sum_{j=1}^{n-k} x_j^2 \sum_{j=1}^{n-k} x_{j+k}^2\right\}^{\frac{1}{2}}}. \tag{1}$$

The figure obtained by plotting r_k as ordinate against k as abscissa and joining each point to the next is known as the correlation diagram or correlogram. Since $r_0 = 1$ the correlogram always starts from the point (0, 1). When the series is infinite in extent, the *serial* correlations become *auto*-correlations (denoted by a Greek ρ), and the series for which such correlations do not vanish may be said to be autocorrelated.

2. If the time series is random the correlogram presents no systematic appearance. If it consists of a simple harmonic the correlogram reproduces that harmonic. If it results from a moving average of finite extent d of a random series the correlogram will, within sampling limits, be zero after $k = d$. The correlogram thus provides a criterion for distinguishing between various kinds of oscillatory time series. In this paper I propose to consider the correlogram of a series defined by the difference equation

$$u_{t+2} + au_{t+1} + bu_t = \epsilon_{t+2}, \tag{2}$$

where a and b are constants and ϵ_{t+2} is a random variable. This equation I shall call the generating equation and the coefficients a and b generating coefficients. The generated series may be said to be autoregressive. The notion of generating series in this way was introduced by Yule (1927) in a classical paper on sunspot periodicities and has been applied to meteorological and economic series with some success. The autoregressive scheme can, in fact, explain a typical phenomenon of such series for which Fourier and periodogram analysis cannot satisfactorily account without great artificiality, namely, the continual shift in phase and variation of amplitude which occur even when the series is smooth. Accounts of the autoregressive scheme have been given in the books by Wold (1938) and Davis (1941).

3. The equation (2) may be solved by the ordinary methods appropriate to difference equations. If the roots of

$$\xi^2 + a\xi + b = 0$$

are $\alpha + i\beta$, $\alpha - i\beta$, the complementary function of (2) is

$$p^t(A\cos\theta t + B\sin\theta t), \tag{3}$$

where $p = +\sqrt{b}$, $\theta = \arctan\frac{\beta}{\alpha} = \arctan\sqrt{\left(\frac{4b}{a^2} - 1\right)}$, and A and B are arbitrary constants.

It is here assumed that b is positive and that $4b > a^2$. It is also to be assumed that $\sqrt{b} = p$ is not greater than unity. The complementary function (3) then represents a damped harmonic. I shall call $2\pi/\theta$ the fundamental period of the generated system.

Let ξ_t be a particular value of (3) such that

$$\xi_0 = 0, \quad \xi_1 = 1, \tag{4}$$

i.e. such that

$$\xi_t = \frac{2}{\sqrt{(4p^2 - a^2)}} p^t \sin \theta t. \tag{5}$$

Then a particular integral of (2) will be found to be

$$\sum_{j=0}^{\infty} \xi_j \epsilon_{t-j+1}. \tag{6}$$

The complete solution is

$$u_t = p^t(A \cos \theta t + B \sin \theta t) + \sum_{j=0}^{\infty} \xi_j \epsilon_{t-j+1}. \tag{7}$$

In practical cases we may assume that the series was 'started up' some time ago, so that the complementary function has been damped out of existence. The series is then given by

$$u_t = \sum_{j=0}^{\infty} \xi_j \epsilon_{t-j+1}. \tag{8}$$

This is a moving sum of a random series with damped harmonic weights. It has been generally assumed, apparently on the basis of experimental evidence, that the mean period of the generated series will be $2\pi/\theta$, the same as the fundamental period present in the term ξ. This, however, is not necessarily so. Something depends on what we call a period in the generated series, whether, for example, we decide to exclude small ripples on the main wave. One possibility is to define the period as the distance between successive 'upcrosses', i.e. points where the series changes sign from negative to positive. The mean distance between major peaks or major troughs will not, probably, be very different from this in the majority of practical cases.*

Some idea of the mean length of the period as so defined can be obtained for particular distributions of ϵ, though a general discussion presents great difficulties. Consider the sum, which will be required again later,

$$\begin{aligned}\sum_{j=0}^{\infty} \xi_j \xi_{j+k} &= \frac{4}{4p^2 - a^2} \Sigma\{p^{2j+k} \sin \theta j \sin \theta(j+k)\} = \frac{2p^k}{4p^2 - a^2} \Sigma[p^{2j}\{\cos \theta k - \cos \theta(2j+k)\}] \\ &= \frac{2p^k}{4p^2 - a^2} \left\{\frac{\cos \theta k}{1 - p^2} - \frac{\cos \theta k - p^2 \cos \theta(k-2)}{1 - 2p^2 \cos 2\theta + p^4}\right\}.\end{aligned} \tag{9}$$

We have

$$\Sigma \xi_j \xi_{j+1} / \Sigma \xi_j^2 = \frac{2p \cos \theta}{1 + p^2} = \frac{-a}{1+b} = \cos \phi, \text{ say.} \tag{10}$$

Now if ϵ is normally distributed, the mean period (from one upcross to the next) is $2\pi/\phi$ (Dodd, 1939). Also it is easily seen that $\cos \theta = \dfrac{-a}{2p} = \dfrac{-a}{2\sqrt{b}}$. Thus the period given by the

* Dodd (1939) points out that in series generated by moving averages of random series there sometimes occur oscillations above or below the x-axis which would not be taken into account by counting upcrosses. The difficulty is to decide which of these is not to be ignored as a 'ripple'. I think that for harmonic weights such as are given by the autoregressive scheme the method of upcrosses is satisfactory, as indicated in paragraph 4.

generating equation is not in general the same as the period in the generated series, defined as the mean difference between upcrosses. The ratio of the first to the second is

$$\frac{\arccos \dfrac{a}{1+b}}{\arccos \dfrac{a}{2\sqrt{b}}}. \tag{11}$$

4. One would expect differences of a similar kind when the variation of ϵ is not normal. Experiments indicate, for example, that rectangular variation of ϵ gives much the same periods as normal variation. It does not seem to have been previously remarked that the mean period of the generated series is not that of the fundamental, but part of the explanation is no doubt due to the fact that for ranges of b encountered in practice (say $0\cdot5 \leqslant b \leqslant 1\cdot0$) the value of the ratio (11) is not very different from unity. In the extreme case $b = 1$ (when the series becomes undamped) the ratio is exactly unity. If $b = 0\cdot5$, $a = -1$, the ratio is 1·07. Notwithstanding the theoretical difference exhibited by equation (11), I think we may take it that for most practical purposes a good estimate of the observed period (upcross to upcross) is that given by the generated equation, *if it is known*. Below I give examples of two artificial series of the type of equation (2) which support this conclusion.

5. Equation (9) provides one further interesting item of information, namely, the relationship between the variance of ϵ and the variance of the generated series. We have

$$\operatorname{var} u = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left(\sum_{j=0}^{\infty} \xi_j \epsilon_{t-j+1} \right)^2 = \sum_{j=0}^{\infty} \xi_j^2 \operatorname{var} \epsilon,$$

cross-product terms in ϵ vanishing since it is a random variable:

$$\frac{\operatorname{var} u}{\operatorname{var} \epsilon} = \Sigma \xi_j^2 = \frac{2}{4p^2-a^2}\left\{\frac{1}{1-p^2} - \frac{1-p^2\cos 2\theta}{1-2p^2\cos 2\theta + p^4}\right\} = \frac{1+b}{(1-b)\{(1+b)^2-a^2\}}.$$

Thus
$$\frac{\operatorname{var} \epsilon}{\operatorname{var} u} = \frac{1-b}{1+b}\{(1+b)^2-a^2\}. \tag{12}$$

In an actual series of 65 terms (referred to in detail below) with $a = -1\cdot1$, $b = 0\cdot5$, the observed ratio of variances was 0·43. The value given by (12) is 0·35. In another series of the same length for which $a = -1\cdot5$, $b = 0\cdot9$, the observed ratio was 0·04, the value given by (12) was 0·07. It is difficult to judge how good the agreement is, but to me it appears satisfactory for such short series. It is noticeable that the generated series may have a very much larger variance than the random series on which it is based.

6. Consider now the autocorrelations of the generated series. We have

$$\begin{aligned}\operatorname{cov}(u_j u_{j+k}) &= \lim \frac{1}{n} \sum_t \{\sum_j (\xi_j \epsilon_{t-j+1}) \sum_j (\xi_j \epsilon_{t+k-j+1})\} = \operatorname{var}\epsilon \sum_{j=0}^{\infty} \xi_j \xi_{j+k} \\ &= \operatorname{var}\epsilon \frac{2p^k}{4p^2-a^2}\left\{\frac{\cos\theta k}{1-p^2} - \frac{\cos\theta k - p^2\cos\theta(k-2)}{1-2p^2\cos 2\theta+p^4}\right\}.\end{aligned}$$

Hence
$$\begin{aligned}\rho_k &= \frac{p^k\{\cos\theta k + \cos\theta(k-2) - 2\cos 2\theta \cos k\theta + p^2\cos\theta k - p^2\cos\theta(k-2)\}}{(1+p^2)(1-\cos 2\theta)} \\ &= \frac{p^k\{\sin(k+1)\theta - p^2\sin(k-1)\theta\}}{(1+p^2)\sin\theta}.\end{aligned}$$

Writing $$\tan\psi = \frac{1+p^2}{1-p^2}\tan\theta,$$

we have $$\rho_k = \frac{p^k(1-2p^2\cos 2\theta+p^4)^{\frac{1}{2}}}{(1+p^2)\sin\theta}\sin(k\theta+\psi) = p^k\frac{\sin(k\theta+\psi)}{\sin\psi}. \tag{13}$$

Apart from the constant factor, ρ_k is thus the product of the damping factor p^k and a harmonic term which has the fundamental period of the generating equation. It is noteworthy that, although the point $\rho_0 = 1$ is a peak at the beginning of the correlogram the existence of the phase angle ψ implies that the interval from $k = 0$ to the next maximum of the correlogram is not equal to the fundamental period. In judging the length of the period from the correlogram it is therefore better to measure from upcross to upcross or from trough to trough; or, if peaks are preferred, not to count the maximum at $k = 0$ as a peak.

The same result may be obtained in two other ways. Multiplying equation (2) by u_{t-k} and summing for all t we have
$$\rho_{k+2}+a\rho_{k+1}+b\rho_k = \frac{\Sigma(\epsilon_{t+2}u_{t-k})}{\operatorname{var} u}.$$

Since u_{t-k} depends only on ϵ_{t-k} and terms with lower subscripts, the expression on the right vanishes if k is not less than -1. We then have
$$\rho_{k+2}+a\rho_{k+1}+b\rho_k = 0 \quad (k \geqslant -1). \tag{14}$$

This result is due to Walker (1931). It was pointed out by Wold that if we multiply (2) by u_{t+k+2} and sum we get
$$\rho_k+a\rho_{k+1}+b\rho_{k+2} = \frac{\Sigma(\epsilon_{t+2}u_{t+k+2})}{\operatorname{var} u}$$

The expression on the right no longer vanishes. In fact u_{t+k+2} contains the term $\xi_{k+1}\epsilon_{t+2}$ and we have
$$\rho_k+a\rho_{k+1}+b\rho_{k+2} = \frac{\operatorname{var}\epsilon}{\operatorname{var} u}\xi_{k+1} \quad (k \geqslant -1). \tag{15}$$

If $k = -1$ the two equations become identical, for $\rho_{-1} = \rho_1$ and we have
$$\rho_1(1+b)+a = 0, \quad \rho_1 = -\frac{a}{1+b}. \tag{16}$$

If we now solve either of the difference equations (14) and (15), making use of the initial conditions $\rho_0 = 1$ and (16) we arrive back at equation (13).

7. The foregoing result (13) would lead us to suppose that the correlogram of an autoregressive series would be damped according to the factor p^k, and this is true for an infinite series. In a number of practical cases, however, I was puzzled by the fact that correlograms of series which appeared on the face of it to be of type (2) did not damp out in the required way. Figs. 1 and 2 show the correlograms of two series for wheat prices and sheep population. The original series were taken from the agricultural returns for England and Wales, trend eliminated by a nine-years moving average and the first thirty serial correlations computed for the resulting series of 64 or 65 terms. The data and the correlations for what prices are given in Tables 1 and 2; those for sheep have been given in a previous paper (Kendall, 1941). I found a similar effect as regards non-damping in nine other agricultural time series, though the correlograms were not so regular as in these two cases. It was always possible, however, that the failure of the fluctuations to damp out in the correlogram was due wholly or partly

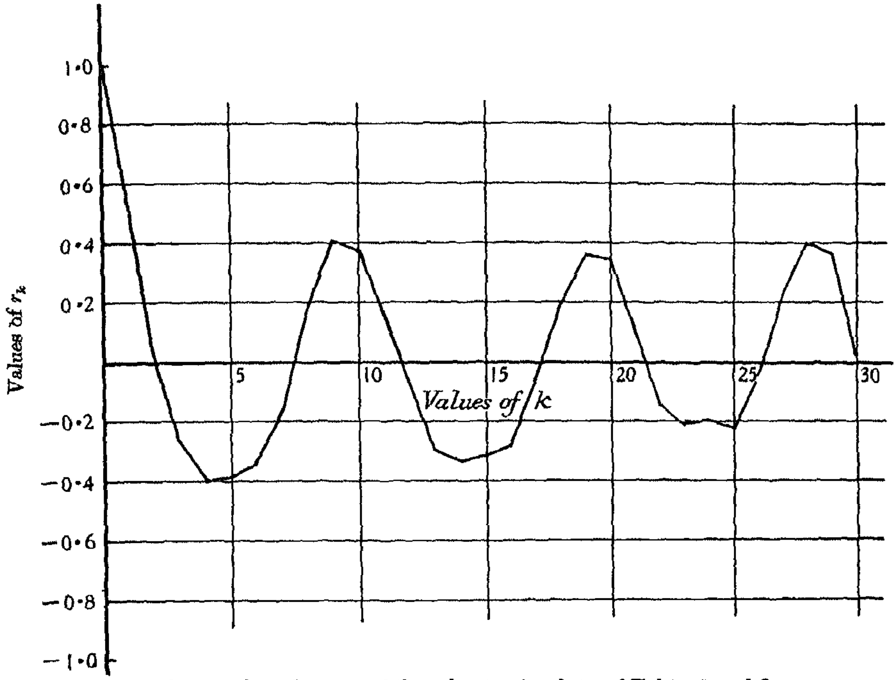


Fig. 1. Correlogram of the wheat price data of Tables 1 and 2.

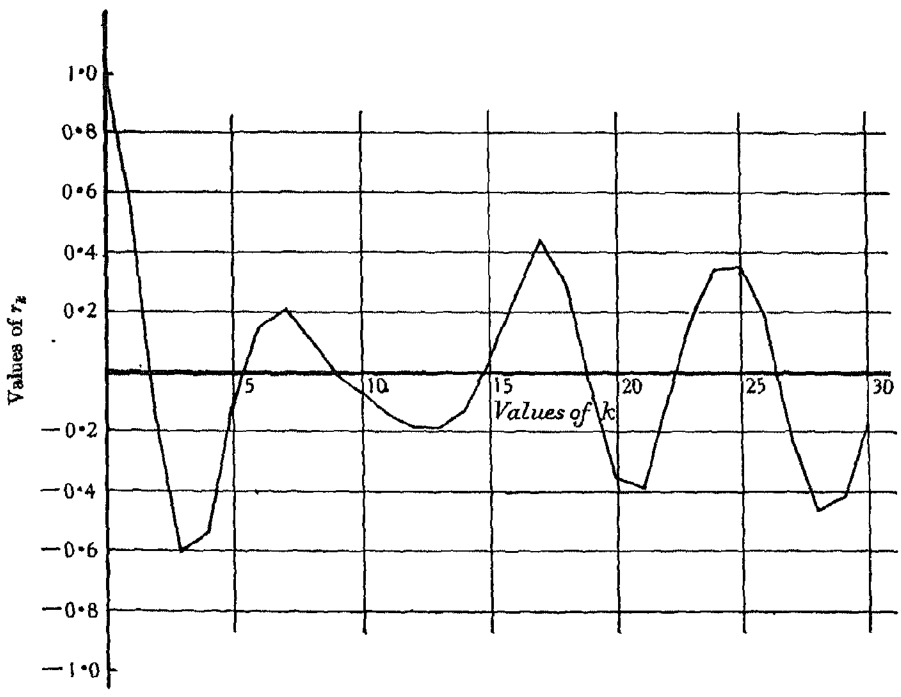


Fig. 2. Correlogram of the sheep population data referred to in paragraph 7.

to failure on the part of the original series to conform to the scheme of equation (2). To avoid *such complications I constructed an artificial series with equation*

$$u_{t+2} = 1{\cdot}1u_{t+1} - 0{\cdot}5u_t + \epsilon_{t+2}. \tag{17}$$

The ϵ's were taken from tables of random numbers and consisted of positive or negative numbers ranging by units from $-9{\cdot}5$ to $+9{\cdot}5$. The series was 'started up' from zero by

Table 1. *Wheat prices,* 1871–1934 *inclusive, deviations from nine-years moving average in pence per hundredweight*

Year	Price	Year	Price	Year	Price
1871	−5	1893	6	1915	−3
2	−13	4	18	6	−7
3	−21	5	15	7	−40
4	−6	6	1	8	−28
5	19	7	−11	9	−27
6	11	8	−17	1920	−51
7	−20	9	5	1	−14
8	3	1900	3	2	37
9	8	1	4	3	42
1880	2	2	2	4	11
1	−6	3	3	5	−9
2	−11	4	−1	6	−17
3	−7	5	2	7	−15
4	6	6	5	8	−5
5	9	7	0	9	−10
6	9	8	−3	1930	−2
7	4	9	−9	1	19
8	3	1910	1	2	14
9	4	1	9	3	21
1890	−3	2	11	4	19
1	−20	3	32		
2	−5	4	33		

Table 2. *Serial correlations of the wheat price series of Table* 1

Order of correlation	r	Order of correlation	r	Order of correlation	r
1	+0·577	11	+0·171	21	+0·115
2	+0·025	12	−0·075	22	−0·144
3	−0·267	13	−0·302	23	−0·204
4	−0·402	14	−0·332	24	−0·195
5	−0·389	15	−0·317	25	−0·221
6	−0·340	16	−0·282	26	−0·052
7	−0·174	17	−0·049	27	+0·217
8	+0·166	18	+0·200	28	+0·404
9	+0·411	19	+0·360	29	+0·364
10	+0·372	20	+0·343	30	+0·024

assuming $u_0 = u_{-1} = 0$. The effect of assuming two consecutive terms equal to zero is not important, for the contribution to a term of the series of terms far back is very small. We are therefore entitled to expect that an artificial series constructed in this way should conform to the foregoing theory. The resultant series and the serial correlations are given in Tables 3 and 4 and the correlogram in Fig. 3. Now this series is heavily damped, p being $\sqrt{(0{\cdot}5)} = 0{\cdot}7071$.

At the twentieth serial correlation, according to equation (14), r_{20} should be less than 0·002 in absolute magnitude. Actually it is one hundred times as big.

8. The explanation lies, I think, in the shortness of the series. In arriving at equation (13) it was assumed that product sums such as $\Sigma \epsilon_j \epsilon_{j+k}$ were zero; and this is so for a series of infinite length. But when the series is short these sums may differ quite appreciably from

Table 3. *Artificial series* $u_{t+2} = 1{\cdot}1u_{t+1} - 0{\cdot}5u_t + \epsilon_{t+2}$ *constructed as described in the text*

No. of term	Value of series	No. of term	Value of series	No. of term	Value of series
1	7	23	−4	45	−13
2	6	24	−5	46	1
3	−6	25	−9	47	6
4	−4	26	−4	48	4
5	3	27	−4	49	11
6	−4	28	3	50	15
7	−5	29	9	51	9
8	−1	30	4	52	8
9	10	31	−8	53	4
10	10	32	−6	54	−1
11	6	33	−3	55	4
12	−4	34	−2	56	7
13	−4	35	0	57	11
14	−7	36	−1	58	0
15	−2	37	−3	59	1
16	6	38	3	60	0
17	17	39	−1	61	−5
18	24	40	−8	62	−11
19	17	41	−3	63	−8
20	4	42	−8	64	−3
21	1	43	−10	65	5
22	−5	44	−16		

Table 4. *Serial correlations of the series of Table* 3

Order of correlation	r	Order of correlation	r	Order of correlation	r
1	+0·70	11	−0·05	21	+0·05
2	+0·29	12	−0·17	22	−0·12
3	+0·01	13	−0·27	23	−0·28
4	−0·17	14	−0·31	24	−0·43
5	−0·27	15	−0·30	25	−0·57
6	−0·25	16	−0·18	26	−0·56
7	−0·13	17	+0·12	27	−0·26
8	+0·07	18	+0·29	28	+0·02
9	+0·12	19	+0·33	29	+0·17
10	+0·05	20	+0·22	30	+0·27

zero. For instance, in samples from a normal population with zero correlation, the standard error of the correlation in samples of 65 is about 0·125, so that values of 0·25 are not very improbable and even values as high as 0·375 are not impossible.

Consider then the series of n terms such as

$$u_j = \xi_1 \epsilon_j + \xi_2 \epsilon_{j-1} + \xi_3 \epsilon_{j-2} + \dots.$$

The product moment of u_j and u_{j+k} will be

$$\sum_j u_j u_{j+k} = \sum_j \sum_l \xi_l \epsilon_{j-l+1} \sum_m \xi_m \epsilon_{j+k-m+1} = \sum_j \sum_{l,m} \xi_l \xi_m \epsilon_{j-l+1} \epsilon_{j+k-m+1}.$$

The terms in ϵ^2 give the expression (13). The others will be sums of products of the ξ's multiplied by the serial covariances of the ϵ's themselves. The dominating terms in these latter will be those containing the larger ξ's. These terms are themselves damped harmonics of the fundamental type and when applied to the sums $\Sigma \epsilon_j \epsilon_{j+k}$ may be expected to generate oscillatory movements of about the same period as the original series. We may therefore expect that for short series the correlogram of the autoregressive system may not decay very rapidly, but that the product terms may themselves result in a small fluctuation. This appears to be happening in both the practical and the artificial examples given above.

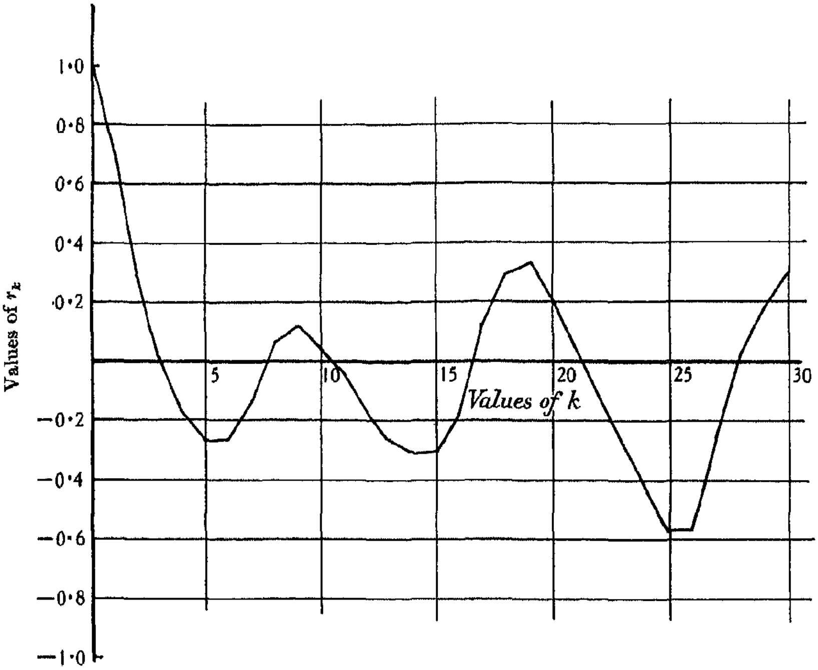


Fig. 3. Correlogram of the artificial series of Tables 3 and 4.

9. Putting aside the mathematics for a moment and looking at the point in a general way, we can, I think, appreciate that something of the kind is to be expected in short series. The variance of a short series should not differ systematically from the variance of the infinite series; but the covariances may systematically exceed in absolute value the value for the infinite series. In fact, as we proceed along the series, the oscillations change in phase, and when we have gone far enough will be quite unrelated in phase to the initial oscillations; but if we only go so far as the second or third oscillation the last oscillation may not, so to speak, have had time to get very much out of phase with the first. The consequence will be that the correlations for such a short space of time will tend to be higher than those for the series as a whole. For the generated system (and indeed for time series in general) we have to be careful not to assume too lightly that values calculated for a part of the series are typical of the corresponding values for the series as a whole; and this notwithstanding that the part of the series was chosen 'at random'.

One would like to know how long a practical series must be for the damping to show itself decisively. An exact discussion of this point presents difficulty, because in a finite series the serial covariances $\Sigma\epsilon_j\epsilon_{j+k}$ are not only non-vanishing but are correlated among themselves. Suppose, however, that we have a series with a damping factor p. Then the kth serial coefficient will not exceed p^k in absolute value, and the difference of the correlogram as a whole from expectation will not at any point exceed the error due to the sampling fluctuation of the serial correlation between u_j and u_{j+k}. This in turn will not exceed the sampling error of the correlation in samples of n from an uncorrelated series, which for large samples has a standard error $n^{-\frac{1}{2}}$. For example, if we took a series of 650 terms instead of 65, correlations up to the 30th would have a standard error not exceeding $1/\sqrt{(650-30)} = 0{\cdot}04$, and hence fluctuations of 0·10 would not be impossible. But by the 30th term in the correlogram the serial correlation has been practically damped out of existence. The fair inference is, I think, that except for the first few serial correlations of the main series, which are not damped very much, the serial correlations are seriously affected even for long series, for the sampling errors are not negligible compared with the small damped 'true' values. In general one would expect the damping effect to show itself for the first five, ten or twenty terms and then to be obliterated by the sampling effect. Whether this happens at the fifth, tenth or twentieth term depends not only on ϵ but on p, the rapidity of damping. In most of the natural series I have examined the damping is fairly rapid, so that the damping effect in the correlogram disappeared after the first few terms. The economic examples given by Wold (1938) and the meteorological examples of Walker (1931) appear to me to support this conclusion.

10. I turn now to consider another effect which may obscure the presence of the generated system of type (2) and may exert an important influence on the correlogram. The random element ϵ considered up to this point is what Yule calls a disturbance, and is integrated into the course of the series by the autoregressive scheme. There may also be a component η *superposed* on the system. This superposed element, if random, is like an error of observation in that its value at any point is unrelated to its value at other points.

If a random element with variance $\operatorname{var}\eta$ is superposed on an infinite series with variance $\operatorname{var}u$ the variance of the whole will be $\operatorname{var}\eta + \operatorname{var}u$. The autocovariances, however, will not be affected (except by sampling effects for short series). Consequently all the autocorrelations except ρ_0 will be reduced in the ratio

$$c = \frac{\operatorname{var}u}{\operatorname{var}u + \operatorname{var}\eta}. \tag{18}$$

For short series there will still be a reduction of the same type, but the value of c may differ from its theoretical value for sampling reasons.

11. An autoregressive series of 65 terms was constructed according to the formula

$$u_{t+2} = 1{\cdot}5u_{t+1} - 0{\cdot}9u_t + \epsilon_{t+2}, \tag{19}$$

where the ϵ_{t+2} were the values of random numbers proceeding by units from $-49{\cdot}5$ to $49{\cdot}5$ with a theoretical variance of 833. On to the series so derived there were superposed (a) a rectangular random element $-49{\cdot}5(1)\,49{\cdot}5$ and (b) a further rectangular random element $-199{\cdot}5(1)\,199{\cdot}5$, additional to the first, the latter series then being divided by ten and rounded up to the nearest integer. The resultant series and serial correlations are shown in Tables 5 and 6, and the correlograms in Fig. 4.

According to (12) the variance of u for an infinite series generated by (19) should be $13{\cdot}97 \operatorname{var}\epsilon$. The values of c for the two series here considered are then, respectively,

13·97/14·963 = 0·93 and 13·97/30·963 = 0·45. In the second case the effect of the superposed variation is to halve the observed correlations. The effect on the actual series of 65 terms is somewhat irregular (see para. 14 below).

Table 5. *Artificial series* $u_{t+2} = 1{\cdot}5u_{t+1} - 0{\cdot}9u_t + \epsilon_{t+2}$, *(a) with small superposed element* η, *(b) with large element* η, *constructed as described in the text*

No. of term	Value of series		No. of term	Value of series		No. of term	Value of series	
	(a)	(b)		(a)	(b)		(a)	(b)
1	5	16	23	−215	−34	45	88	3
2	36	7	24	−219	−15	46	76	− 5
3	8	10	25	− 76	−17	47	93	21
4	− 81	− 5	26	95	− 4	48	34	8
5	− 89	7	27	239	15	49	60	15
6	− 15	− 7	28	316	23	50	− 69	−13
7	35	10	29	289	21	51	−120	−10
8	112	13	30	169	13	52	− 54	12
9	146	17	31	49	− 7	53	− 56	−13
10	100	19	32	−114	−26	54	12	7
11	1	7	33	−259	−26	55	5	−11
12	−131	1	34	−290	−26	56	13	−18
13	−195	− 6	35	−208	−29	57	3	− 1
14	−259	−27	36	− 31	− 4	58	13	11
15	−258	−37	37	21	12	59	− 4	9
16	−118	−28	38	109	18	60	− 71	−13
17	32	6	39	26	19	61	− 94	− 7
18	110	− 3	40	33	− 5	62	15	19
19	245	6	41	22	0	63	45	− 3
20	166	24	42	− 30	−10	64	138	19
21	79	15	43	51	−15	65	115	0
22	−177	−34	44	49	17			

Table 6. *Serial correlations of series (a) and (b) of Table 5*

Order of correlation	r		Order of correlation	r		Order of correlation	r	
	(a)	(b)		(a)	(b)		(a)	(b)
1	+0·78	+0·49	11	+0·33	+0·16	21	+0·21	+0·05
2	+0·33	+0·13	12	−0·04	−0·16	22	−0·07	−0·28
3	−0·22	−0·13	13	−0·38	−0·37	23	−0·28	−0·33
4	−0·63	−0·42	14	−0·58	−0·50	24	−0·32	−0·29
5	−0·75	−0·46	15	−0·52	−0·36	25	−0·22	−0·20
6	−0·59	−0·39	16	−0·23	−0·14	26	−0·01	−0·07
7	−0·22	−0·01	17	+0·14	+0·19	27	+0·16	+0·11
8	+0·19	+0·28	18	+0·43	+0·42	28	+0·23	+0·20
9	+0·48	+0·38	19	+0·55	+0·41	29	+0·13	+0·01
10	+0·53	+0·52	20	+0·45	+0·37	30	−0·11	−0·11

The correlograms run according to expectation. The effect of the bigger random element is to reduce the amplitude at the beginning of the series and to introduce some minor irregularities in the data, but not to affect substantially the lengths of the correlogram oscillations.

12. But here arises one important difficulty. Suppose we are given such series as these and require to estimate a and b, the constants of the generating equation. The procedure adopted by Yule was as follows: we take the observed serial correlations r_1 and r_2 as estimates of the autocorrelations. We then find the regression equation of u_{t+2} on u_{t+1} and u_t by the usual methods, assuming that the variance of the series is the same as the variance for the series less its first term and that the serial correlation r_1 for the whole series is the same as that for the series less its first term. The regression equation is

$$u_{t+2} = \frac{r_1(1-r_2)}{1-r_1^2} u_{t+1} + \frac{r_2 - r_1^2}{1-r_1^2} u_t. \tag{20}$$

The observed regression equation is then taken as an estimate of the generating equation so that we have as estimates of a and b

$$-a = \frac{r_1(1-r_2)}{1-r_1^2}, \quad -b = \frac{r_2 - r_1^2}{1-r_1^2} = \frac{r_2 - 1}{1-r_1^2} + 1. \tag{21}$$

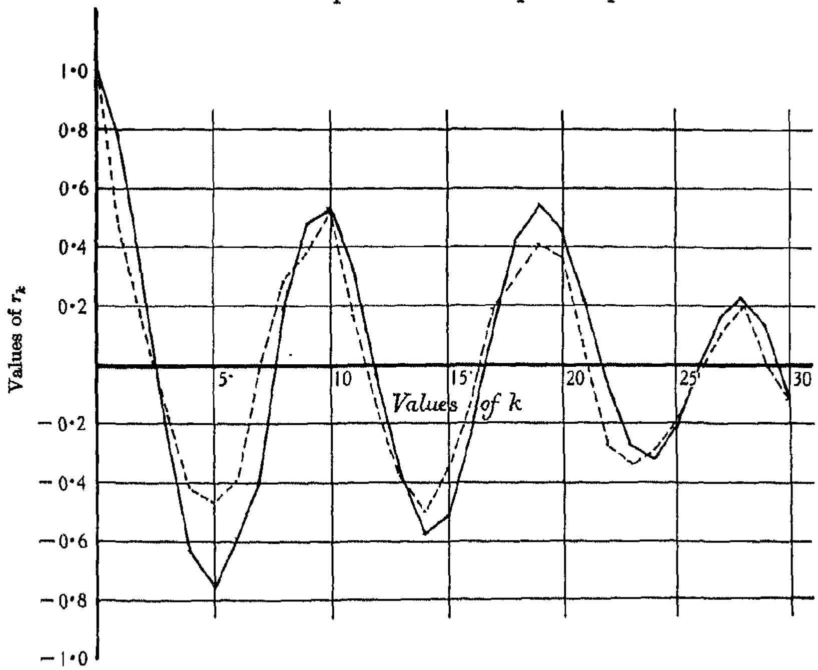


Fig. 4. Correlogram of the two artificial series of Tables 5 and 6, the full line representing series (a) with slight superposed variation, the broken line series (b) with large superposed variation.

From the above it will be evident that there are two sources of possible error in the use of these equations: (1) if the series is short r_1 and r_2 may not be reliable estimates of the autocorrelations for the infinite series, (2) if there is any superposed variation the observed r's will be lower than the true r's for the autoregressive system, being in fact cr_1 and cr_2 where c is given by (18). Consider the second of these effects.

13. In the case of superposed variation the use of (21) will lead to the equations

$$-a' = \frac{cr_1(1-cr_2)}{1-c^2r_1^2}, \quad -b' = \frac{cr_2 - c^2r_1^2}{1-c^2r_1^2}.$$

The estimated fundamental period of the generating equation is then given by

$$4\cos^2\theta' = \frac{a'^2}{b'} = \frac{cr_1^2(1-cr_2)^2}{(1-c^2r_1^2)(r_2-cr_1^2)}. \tag{22}$$

If we expand (22) in powers of $\gamma = 1-c$ we find, to first order in γ,

$$\frac{a'^2}{b'} = \frac{a^2}{b}\left\{1-\gamma\frac{(1+b)(3b^2-b-a^2)}{b\{(1+b)^2-a^2\}}\right\}.$$

Hence, if $3b^2-b-a^2>0$ the effect of a superposed variation (equivalent to positive γ) is to give

$$\frac{a'^2}{b'}<\frac{a^2}{b},$$

or in other words to result in a shortening of the observed period. The condition that $3b^2-b-a^2>0$ is equivalent to

$$b>\tfrac{1}{6}\{-1+\sqrt{(12a^2+1)}\}$$

and is not very restrictive since in any case $a^2 \leqslant 4$ and $4b \geqslant a^2$. The inequality is obeyed by all the examples I have met in practice.

We therefore reach the interesting conclusion that if there is any superposed random variation present, the period calculated from the observed regression equation according to formulae (21) will probably be too short even for long series. Yule himself found too short a period for his sunspot material and, suspecting that it was due to superposed variation, attempted to reduce that variation by graduation. The result was to give a longer period more in accordance with observation. It does not appear, however, that the superposed variation in his case was very big. In a number of agricultural time series which I have examined it is sometimes about half the variation of the series and the effect on the period as calculated from the serial correlations is very serious. For instance, in the cases of wheat prices and sheep population referred to above, formulae (21) give periods of 7·0 and 6·8 years, whereas the correlograms indicate periods of about 9·5 and 8·5 years respectively.

14. As an example, consider the second of the artificial series referred to in paragraph 11. For the observed serial correlations I find

$$r_1' = 0{\cdot}486, \quad r_2' = 0{\cdot}133,$$

giving, according to (21),

$$-a' = 0{\cdot}552, \quad b' = 0{\cdot}135, \quad \cos\theta' = \frac{-a'}{2\sqrt{b'}} = 0{\cdot}751, \quad \theta' = 41{\cdot}3^\circ,$$

which corresponds to a period of about 8·7 years, whereas we know from the construction of the series that

$$a = 1{\cdot}5, \quad b = 0{\cdot}9,$$

giving $\theta = 37{\cdot}7^\circ$ and a period of 9·5 years.

Considering the profound effect which the superposed variation has had on the first two serial coefficients, reducing r_1 from 0·78 to 0·49 and r_2 from 0·33 to 0·13, one might have expected the period to have been affected even more than appears from this result. But the example serves to bring out once again the difficulties associated with short series and the unreliability of coefficients calculated from the first two serial correlations in such cases.

If, for instance, we had found $r_2' = 0{\cdot}18$ instead of $0{\cdot}13$ we should have obtained a period of about 12 years, and if $r_2' \doteq 0{\cdot}20$ the solution becomes impossible, for a' and b' then assume values such that $a'^2 > 4b'$ and $\cos\theta' > 1$. Furthermore, such a value as $0{\cdot}18$ for r_2' increases the period instead of decreasing it. We note, in fact, that r_2 has been reduced by $0{\cdot}13/0{\cdot}33 = 0{\cdot}40$ so that it is not legitimate to assume that r_1 and r_2 are reduced by a constant c. The errors introduced by neglect of the autocovariances of the superposed random element may be so serious as to destroy the value of calculations based on the observed regression coefficients.

15. Even in long series, when it is legitimate to suppose that r_1 and r_2 are reduced in the same proportion, the length of the period is very sensitive to superposed variation. Consider, for example, the effect of small variations in c near $c = 1$. We have from (22), differentiating logarithmically and putting $c = 1$,

$$-2\tan\theta\frac{d\theta}{dc} = 1 - \frac{2r_2}{1-r_2} + \frac{2r_1^2}{1-r_1^2} + \frac{r_1^2}{r_2 - r_1^2},$$

which on substituting $$r_1 = -\frac{a}{1+b}, \quad r_2 = -b + \frac{a^2}{1+b}$$

reduces to $$-\tan\theta\frac{d\theta}{dc} = \frac{(1+b)(3b^2+b-a^2)}{2b\{(1+b)^2-a^2\}}. \tag{23}$$

Now the period $$P = 2\pi/\theta \quad \text{and} \quad \tan\theta = \sqrt{\left(\frac{4b}{a^2}-1\right)}.$$

Hence $$\frac{dP}{dc} = -\frac{P^2 a(1+b)(3b^2+b-a^2)}{4\pi\sqrt{(4b-a^2)}\{(1+b)^2-a^2\}}. \tag{24}$$

When $a = -1{\cdot}5$, $b = 0{\cdot}9$, $P = 9{\cdot}5$, this reduces to

$$\frac{dP}{dc} = 14 \text{ intervals.}$$

Thus if $c = 0{\cdot}9$, i.e. the superposed variation is only about 10 % of the total, the period may be shortened by something of the order of one interval.

16. The position may then be summarized as follows:

(*a*) The correlogram of a generated system of the type of equation (2) will be damped according to the damping factor of the equation; but if the series is short the damping may be considerably less than the theoretical value.

(*b*) The correlogram will show a period equal, within limits of error, to that of the fundamental period of the system. The distance from the unit ordinate at $k = 0$ and the first maximum of the correlogram may not, however, be a full period and in estimating periods from the correlogram it is better to reckon from upcross to upcross.

(*c*) It does not appear that in short series the periodic movement is substantially affected, at any rate not to such an extent as the damping.

(*d*) When superposed random variation is present the fundamental period calculated from the observed regressions will be too short and may be very considerably so.

(*e*) The period of the generated series, defined as the mean distance from upcross to upcross, is not quite the same as the fundamental period: but the difference is not likely to be important in practice.

17. Armed with these results we may consider the difficult problem of determining, for a given time series, the autoregressive scheme which may have generated it.

The first step is to calculate the serial correlations and examine the correlogram. If the latter shows fairly regular oscillations there is a presumption that the series is of the autoregressive type, and this conclusion need not be rejected solely because the oscillations do not damp so rapidly in the later portion of the correlogram as at the beginning. To take the wheat price material given above, an examination of the correlogram (Fig. 1) strongly suggests a simple damped harmonic. The sheep series (Fig. 2) is not so clearly defined and there is a suggestion of more than one period in the behaviour of the correlogram. This may, however, be due to the shortness of the series and one would be inclined to consider it in the first place as a single damped harmonic.

The length of the period, as pointed out above, is best determined by the mean distances between upcrosses. In the wheat price data there are upcrosses at about 7·5 years, 17·2 years and 26·1 years, giving periods of 9·7 and 8·9 years with a mean of 9·3 years. Much the same result is arrived at by counting the periods between troughs.

18. The second step is to calculate a' and b' by equations (21) and to find the period of the fundamental harmonic as given by the observed regression equation. For wheat prices we have

$$r'_1 = 0{\cdot}5773, \quad r'_2 = 0{\cdot}0246.$$

Hence

$$a' = -0{\cdot}8446, \quad b' = 0{\cdot}4630,$$

whence

$$\cos\theta' = 0{\cdot}6206, \quad \theta' = 51{\cdot}63^\circ,$$

giving a fundamental period of 6·97 years.

This is too small, and we are immediately led to suspect the existence of superposed variation. The problem then arises of determining the variance of the superposed element. If this is random, and there are no periodic terms of very short period the variate difference method may be used. For the wheat price material, taking differences up to the 10th on the primary series (before trend was eliminated) I find for an estimate of the random variance $\operatorname{var}\eta = 27{\cdot}72$. The total variance of the series is 272·8. The constant c of equation (18) is

$$1 - \frac{27{\cdot}72}{272{\cdot}8} = 0{\cdot}90.$$

The putative serial correlations of the autoregressive series will then be

$$r_1 = 0{\cdot}641, \quad r_2 = 0{\cdot}027$$

and

$$a = -1{\cdot}059, \quad b = 0{\cdot}652,$$

whence

$$\cos\theta = 0{\cdot}6551, \quad \theta = 49{\cdot}07^\circ,$$

giving a period of 7·34 years.

This is still too short. It would require a random superposed variance of about 25 % of the total, instead of the observed 10 %, to produce a period of 9·3 years.

19. This illustrates very well a constantly recurring difficulty in the theory of the variate difference method and of time series generally. The superposed variation may not be random. Indeed, we have little ground for expecting that it should be. A positive correlation between the successive values of η will reduce the variance shown as random by the variate difference method and unless we have prior reason to suppose that η is random the values given by the variate difference method are quite likely to be too small. Unfortunately we rarely have any prior knowledge of η, but from general economic considerations one would not be surprised

to find that there do exist positive correlations from one year to the next, owing to the enduring nature of some of the causes which can give rise to superposed variation. I conclude generally that discrepancies of the type here considered support the view that the period is to be determined from the correlogram, not from solution of the regression equation.

20. Two points may be mentioned incidentally. One, a matter of technique, is that the arithmetic of serial correlations and variate differences are closely linked together and the results of the one can be used to derive those of the other. This means an enormous saving in arithmetic and the method is described in the Appendix to this paper.

The second point concerns the removal of superposed random variation by graduation formulae. It will be clear that if the superposed variation is not random graduation may only make matters worse; but even if it is random, graduation formulae may induce spurious cyclical effects into the data. It seems to me that as a general rule graduation is to be undertaken with great caution.

21. Reverting to the main topic, it would seem that if the period shown by the correlogram and the period calculated from the observed regression equation disagree, and cannot be reconciled by the assumption of a superposed random element, there is little further to be done to dissect the superposed element from the autoregressive part of the system. If, however, the variate difference method supplies a variance of η which can satisfactorily explain the difference in periods, we may go forward. The constant c can be calculated and the constants of the autoregressive part of the system determined. Equation (12) then gives an estimate of the ratio between the variance of the basic random element ϵ and the variance of the generated series. If there is a superposed element it will not be possible to find the values of that element at every point and so to determine ϵ at every point; but if η is non-existent, ϵ can be explicitly determined, except for the first two terms of the series. In fact, we merely apply the generating equation to the observed series, the residuals between prediction and observation being the values of ϵ. An examination of these residuals will confirm whether they may be regarded as a random series.

22. The foregoing treatment can be extended to the case of a more general linear regressive system

$$u_{t+m} + a_1 u_{t+m-1} + \dots + a_m u_t = \epsilon_{t+m},$$

or to cases where the regression is curvilinear; and the necessity for more general schemes in representing observed data can, as shown by Yule, be discussed in terms of partial autocorrelations and scatter diagrams. A more serious problem arises if the series ϵ is itself not random, a state of affairs which one fears might be fairly common in economic series. To take the wheat price data once again, it would not be surprising to find that the wheat price oscillations were regenerated by a series of disturbances, part of which were attributable to variations in acreages, yields, or the prices of other crops. Such disturbances might themselves be oscillatory. For such cases the problem becomes exceedingly complicated. To discuss it at all satisfactorily one would require a long series or collateral evidence in the form of other series of a similar character. If there is a royal road in this subject it has not yet been discovered.

APPENDIX

Relationship between variate differences and serial correlations

If we have a series of values $x_1 \dots x_n$ the first differences are $x_1 - x_2$, etc., the second differences $x_1 - 2x_2 + x_3$, etc., and so on. Let S_j be the sum of the squares of the jth differences and write for the product-sums

$$P_j = \sum_{k=1}^{n-j} x_k x_{k+j}. \tag{25}$$

The sums S are those appearing in variate difference analysis and the quantities P appear in serial correlation analysis. Either set can be expressed in terms of members of the other, as follows:

We have
$$S_0 = P_0, \tag{26}$$

$$S_1 = \sum (x_j - x_{j+1})^2 = \sum_{j=1}^{n-1} x_j^2 - 2\sum_{j=1}^{n-1} x_j x_{j+1} + \sum_{j=2}^{n} x_j^2 = 2P_0 - x_1^2 - x_n^2 - 2P_1, \tag{27}$$

$$\begin{aligned} S_2 = \sum (x_j - 2x_{j+1} + x_{j+2})^2 &= \sum_{j=1}^{n-2} x_j^2 + 4\sum_{j=2}^{n-1} x_j^2 + \sum_{j=3}^{n} x_j^2 - 4\sum_{j=1}^{n-2} x_j x_{j+1} + 2\sum_{j=1}^{n-2} x_j x_{j+2} - 4\sum_{j=2}^{n-1} x_j x_{j+1} \\ &= 6P_0 - 8P_1 + 2P_2 - x_1^2 - x_n^2 - (2x_1 - x_2)^2 - (2x_n - x_{n-1})^2 \end{aligned} \tag{28}$$

and so on. For the purpose of expressing the general formulae of this kind it is convenient to modify the sums S. Suppose we write the series $x_1 \dots x_n$ preceded and followed by a number of zeros. The difference table will then appear as follows:

$$\begin{array}{ccccc}
0 & & & & \\
 & 0 & & & \\
0 & & 0 & & \\
 & 0 & & -x_1 & \\
0 & & +x_1 & & \\
 & -x_1 & & 3x_1 - x_2 & \text{etc.} \\
x_1 & & -2x_1 + x_2 & & \\
 & x_1 - x_2 & & -3x_1 + 3x_2 - x_3 & \\
x_2 & & x_1 - 2x_2 + x_3 & & \\
 & x_2 - x_3 & & x_1 - 3x_2 + 3x_3 - x_4 & \\
x_3 & & x_2 - 2x_3 + x_4 & &
\end{array}$$

with a symmetrical effect at the other end. Writing now T_1, T_2, etc., for the sum of the squares of members in the first, second, etc., column of differences we see that

$$T_j = \sum_k \left\{ x_k - \binom{j}{1} x_{k+1} + \binom{j}{2} x_{k+2} \dots \right\}^2, \tag{29}$$

where the summation now takes place over all values of x and there are no complications introduced by end effects. In fact, we have thrown the end effects into the sums T which replace the S's. In actually calculating the T's from the S's it is very little trouble to add the extra terms to the tables giving the latter; and when calculating the S's from the T's only the differences at the end of the table need be worked out.

We have then from (29), on expansion

$$T_j = P_0\left\{\binom{j}{0}^2+\binom{j}{1}^2+\dots\right\}+2(-1)^k\sum_k\left[P_k\left\{\binom{j}{0}\binom{j}{k}+\dots+\binom{j}{j-k}\binom{j}{j}\right\}\right]. \tag{30}$$

The coefficients of the various P's are easily seen to be equal to corresponding powers of t in

$$\left\{\binom{j}{0}t^0-\binom{j}{1}t+\binom{j}{2}t^2\dots\right\}\left\{\binom{j}{0}t^j-\binom{j}{1}t^{j-1}+\dots\right\},$$

i.e. in $(-1)^j(1-t)^{2j}$, and we find, on substitution in (30),

$$T_j = P_0\binom{2j}{j}-2P_1\binom{2j}{j-1}+2P_2\binom{2j}{j-2}+\dots+2(-1)^jP_j. \tag{31}$$

For example

$$\begin{aligned}
T_0 &= P_0,\\
T_1 &= 2P_0-2P,\\
T_2 &= 6P_0-8P_1+2P_2,\\
T_3 &= 20P_0-30P_1+12P_2-2P_3,\\
T_4 &= 70P_0-112P_1+56P_2-16P_3+2P_4,\\
T_5 &= 252P_0-420P_1+240P_2-90P_3+20P_4-2P_5,\\
T_6 &= 924P_0-1584P_1+990P_2-440P_3+132P_4-24P_5+2P_6,\\
T_7 &= 3432P_0-6006P_1+4004P_2-2002P_3+728P_4-182P_5+28P_6-2P_7,\\
T_8 &= 12870P_0-22880P_1+16016P_2-8736P_3+3640P_4-1120P_5+240P_6-32P_7+2P_8,\\
T_9 &= 48620P_0-87516P_1+63648P_2-37128P_3+17136P_4-6120P_5+1632P_6-306P_7\\
&\quad +36P_8-2P_9,\\
T_{10} &= 184756P_0-335920P_1+251940P_2-155040P_3+77520P_4-31008P_5+9690P_6\\
&\quad -2280P_7+380P_8-40P_9+2P_{10}.
\end{aligned} \tag{32}$$

The coefficients check in virtue of the fact that they sum to zero.

Conversely we have

$$\begin{aligned}
2P_0 &= 2T_0,\\
2P_1 &= -T_1+2T_0,\\
2P_2 &= T_2-4T_1+2T_0,\\
2P_3 &= -T_3+6T_2-9T_1+2T_0,\\
2P_4 &= T_4-8T_3+20T_2-16T_1+2T_0,\\
2P_5 &= -T_5+10T_4-35T_3+50T_2-25T_1+2T_0,\\
2P_6 &= T_6-12T_5+54T_4-112T_3+105T_2-36T_1+2T_0,\\
2P_7 &= -T_7+14T_6-77T_5+210T_4-294T_3+196T_2-49T_1+2T_0,\\
2P_8 &= T_8-16T_7+104T_6-352T_5+660T_4-672T_3+336T_2-64T_1+2T_0,\\
2P_9 &= -T_9+18T_8-135T_7+546T_6-1287T_5+1782T_4-1386T_3+540T_2-81T_1+2T_0,\\
2P_{10} &= T_{10}-20T_9+170T_8-800T_7+2275T_6-4004T_5+4290T_4-2640T_3+825T_2\\
&\quad -100T_1+2T_0.
\end{aligned} \tag{33}$$

The coefficients sum in turn to 2, 1, -1, -2, -1, 1, 2, etc. The coefficient of T_k in $2P_j$ is

$$(-1)^{j-k}\frac{2j}{j+k}\binom{j+k}{2k}.$$

REFERENCES

DAVIS, H. T. (1941). *The Analysis of Economic Time Series*. Bloomington, Indiana.

DODD, E. L. (1939). The length of the cycles which result from the graduation of chance elements. *Ann. Math. Statist.* **10**, 254.

KENDALL, M. G. (1941). The effect of the elimination of trend on oscillations in time series. *J. Roy. Statist. Soc.* **104**, 43.

YULE, G. UDNY (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philos. Trans.* A, **226**, 267.

WALKER, SIR GILBERT (1931). On periodicity in series of related terms. *Proc. Roy. Soc.* A, **131**, 518.

WOLD, H. (1938). *A Study in the Analysis of Stationary Time-series*. Uppsala.

COMPARISON OF THE CONCEPTS OF EFFICIENCY AND CLOSENESS FOR CONSISTENT ESTIMATES OF A PARAMETER

By R. C. GEARY

Having given two estimates X and Y of an unknown parameter θ, E. J. G. Pitman (1937) has suggested that X should be regarded as a 'closer' estimate of θ if the probability that

$$|X-\theta|<|Y-\theta| \tag{1}$$

is greater than $\frac{1}{2}$. This concept has intuitive appeal and it is in accordance with statistical tradition that preference should be expressed on a probability scale. The object of this communication is to compare 'closeness', as defined, with the familiar concept of relative 'efficiency' (Fisher, 1922) as determined by the variances of the two estimates. Continuous variation is assumed throughout.

Pitman has shown that the median has a special role in his theory of closeness and, since the median is notably unamenable to algebraic treatment, it is not to be expected that, despite its apparent simplicity, condition (1) should be readily expressible in terms of the semi-invariants (assuming these known firmly or approximately) of the joint distribution of X and Y.

Suppose that X and Y are consistent estimates of θ, i.e. that

$$EX = EY = \theta. \tag{2}$$

In this connexion it must be observed that, in not necessarily large sample theory, this condition need not be observed. In fact, Pitman (1937, p. 215) has shown that in estimating the variance of normal samples of n, the 'closest' estimate of the variance is

$$s_2' = \sum_{i=1}^{n}(x_i-\bar{x})^2/(n-\tfrac{2}{3}), \text{ approximately,} \tag{3}$$

which is not consistent, instead of the usual

$$s_2'' = \sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1), \tag{4}$$

which is.

If the joint distribution of X and Y is symmetrical, i.e. if

$$\lambda_{ij} = \lambda_{ji} \tag{5}$$

for all joint semi-invariants of powers i and j, then it is clear that

$$\text{Prob}\{|X-\theta|<|Y-\theta|\} = \text{Prob}\{|Y-\theta|<|X-\theta|\},$$

and, since the sum of the two probabilities is unity, each must be equal to $\frac{1}{2}$. It follows that, when the two estimates are distributed on the normal surface of error with equal variances, the probability of (1) is $\frac{1}{2}$, since in this case $\lambda_{ij} = \lambda_{ji}$, the semi-invariants being zero for all values of i and j except those for which $(i+j) = 2$, where the sets for (i, j) of $(2, 0)$ and $(0, 2)$

are the variances of X and Y respectively. It can readily be shown that the probability of (1) in the normal case, with variances of σ_x^2 and σ_y^2 and coefficient of correlation ρ is

$$\frac{1}{\pi}\tan^{-1}\left(\frac{2\sigma_x\sigma_y\sqrt{(1-\rho^2)}}{\sigma_x^2-\sigma_y^2}\right), \tag{6}$$

which shows that in this case the criteria of closeness and efficiency are identical, since, for all values of ρ, the probability is greater than, or less than, $\frac{1}{2}$ according as σ_x is respectively less than, or greater than, σ_y.

Under general conditions the joint distribution of

$$x = X-\theta \quad\text{and}\quad y = Y-\theta$$

will be of the form

$$f(x,y) = \sum_{i+j=3}^{\infty}\sum \frac{(-)^{i+j}}{i!\,j!}\lambda_{ij}\left(\frac{\partial}{\partial x}\right)^i\left(\frac{\partial}{\partial y}\right)^j \Phi(x,y)\,dx\,dy, \tag{7}$$

$\Phi(x,y)$ being the normal function

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2}-\frac{2\rho xy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}\right)\right\}\Big/\,2\pi\sqrt{(1-\rho^2)}. \tag{8}$$

If the two estimates are computed from samples of n it is well known that

$$\lambda_{ij}/\sigma_x^i\sigma_y^j$$

is of order $\frac{1}{2}(2-i-j)$ in n. It is natural, accordingly, to try to find the value of probability of (1) when the variates are distributed in (7) assuming that the λ_{ij} are zero for all but the smaller values of $(i+j)$. Probability of (1) is given by

$$\int_{-\infty}^{+\infty}dy\int_{-|y|}^{+|y|}f(x,y)\,dx. \tag{9}$$

To find the value of this integral the frequency $f(x,y)$ is regarded as the product of the normal function Φ and a polynomial in x and y determined from (7). It can easily be shown that polynomials of odd order in $(i+j)$ vanish on integration. Accordingly the integral (9) is given by (6) plus a term $O(n^{-1})$; furthermore the expression for the probability contains only terms in n^{-k}, where k is a positive integer.

There is no theoretical difficulty about the computation of the coefficients of powers of $\lambda_{ij}/\sigma_i\sigma_j$ in (9) but the algebra is complicated. For the purpose of this note it will suffice perhaps to state that when

$$\sigma_x = \sigma_y,\ \lambda_{ij} = 0 \ \text{for}\ (i+j) > 4 \quad\text{and}\quad \lambda_{ij}^k = 0 \ \text{for}\ k>2,\ (i+j) = 3,$$

then the probability (1) is given by

$$\begin{aligned}\frac{1}{2}+\frac{1}{72\pi}(1-\rho^2)^{-\frac{3}{2}}\{6(1-\rho^2)(\lambda'_{40}-\lambda'_{04})-12\rho(1-\rho^2)(\lambda'_{31}-\lambda'_{13})&\\ -(7+2\rho^2)(\lambda'^2_{30}-\lambda'^2_{03})+36\rho(\lambda'_{30}\lambda'_{21}-\lambda'_{12}\lambda'_{03})&\\ -3(1+2\rho^2)(2\lambda'_{30}\lambda'_{12}+3\lambda'^2_{21}-2\lambda'_{21}\lambda'_{03}-3\lambda'^2_{12})\},&\end{aligned} \tag{10}$$

with

$$\lambda'_{ij} = \lambda_{ij}/\sigma_x^i\sigma_y^j.$$

The term in addition to $\frac{1}{2}$ may be regarded as $O(n^{-1})$, the terms neglected being $O(n^{-2})$ when n is large. When n is not large it is evident that the added term may be quite appreciable

when λ'_{ij} is significantly different from λ_{ji} and more particularly if ρ is nearly unity, which will usually be the case with two estimates which are of almost equal efficiency. In a particular example (which, however, had no reference to the problem of estimation) the values found were as follows:

$$\rho = 0{\cdot}953; \quad \lambda'_{30} = -0{\cdot}273; \quad \lambda'_{21} = -0{\cdot}287; \quad \lambda'_{12} = -0{\cdot}302; \quad \lambda'_{03} = -0{\cdot}307;$$
$$\lambda'_{40} = -0{\cdot}563; \quad \lambda'_{31} = -0{\cdot}513; \quad \lambda'_{22} = -0{\cdot}466; \quad \lambda'_{13} = -0{\cdot}418; \quad \lambda'_{04} = -0{\cdot}327.$$

For these values of the semi-invariants the value of (10) is 0·5185, which does not differ much from the value $\frac{1}{2}$ which would be obtained on the assumption of equal variances and normal distribution of the estimates; it is equally certain that, when the estimates are computed from large random samples and when their joint distribution tends towards normality as the sample number tends towards infinity, the probability (1) will be exceedingly close to $\frac{1}{2}$; there is no good *theoretical* reason, however, for thinking that for estimates computed from samples which are not large, the value of (10) (or of the more exact value of the probability which would be found by taking into account further terms in the expansion of (7) in the compilation of (9)) would be close to $\frac{1}{2}$, in general.

SOME PARTICULAR CASES

As an application, consider the method of estimating the universal mean of a normal distribution, give a random sample of n, assuming that the universal variance is unity. The distribution is, accordingly,

$$\frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}(z-\theta)^2} dz.$$

Since the maximum likelihood estimate of θ is the mean

$$X = \frac{1}{n} \sum_{i=1}^{n} z_i$$

its variance in large samples must be less than that of any other estimate Y of the parameter (Fisher, 1922, 1935). It is well-known that the median y of large samples of n from a normal universe with mean zero and variance unity is distributed approximately as

$$\frac{1}{\pi} \sqrt{\left(\frac{n}{2}\right)} e^{-ny^2/\pi} dy,$$

whereas the mean x is distributed as

$$\sqrt{\left(\frac{n}{2\pi}\right)} e^{-nx^2/2} dx,$$

so that the respective variances are $\sqrt{(\pi/2n)}$ and $1/\sqrt{n}$. It has been computed that for large normal samples the correlation between the mean and the median is $\rho = \sqrt{(2/\pi)}$. From formula (6)

$$\text{Prob}\,\{| X - \theta | < | Y - \theta |\} = 0{\cdot}615.$$

Pitman (1937, p. 221) has shown that the closest estimate of the centre of a rectangular distribution is the mid-point (or the mean of the largest and smallest members of the sample). The rest of this section deals with comparisons, by means of the criteria of closeness and efficiency, of the mid-point and other consistent estimates of the centre of the rectangular distribution of which the range is known to be unity.

It is a well-known fact that by the variance test the mid-point is more efficient than the *arithmetic mean* as an estimate of the middle point of the range for a rectangular universe when the range is known. In fact the respective variances are as follows:

	Variance
Mid-point:	$1/\{2(n+1)(n+2)\}$
Mean:	$1/\{12n\}$,

so that the variance of the mid-point is actually of a lower order of magnitude in n than the variance of the mean. The respective distributions (when the centre is zero) are as follows:

Mid-point: $$n(1-2\,|\,t\,|)^{n-1}dt \tag{11}$$

Mean (Gram-Charlier, correct to n^{-3}):

$$\frac{dy}{\sqrt{(2\pi)}}\left[1-\frac{1}{20n}\left(\frac{d}{dy}\right)^4+\frac{1}{n^2}\left\{\frac{1}{105}\left(\frac{d}{dy}\right)^6+\frac{1}{800}\left(\frac{d}{dy}\right)^8\right\}\right.$$
$$\left.-\frac{1}{n^3}\left\{\frac{3}{1{,}400}\left(\frac{d}{dy}\right)^8+\frac{1}{2{,}100}\left(\frac{d}{dy}\right)^{10}+\frac{1}{48{,}000}\left(\frac{d}{dy}\right)^{12}\right\}+\dots\right]e^{-y^2/2}, \tag{12}$$

$y = x\sqrt{(12n)}$, x being the mean of sample of n.

The distributions of the rectangular mid-point and lowest point (used later) have been given by Neyman & Pearson (1928). The β_2 of the distribution of the mid-point is given by

$$\beta_2 = \frac{6(n+1)(n+2)}{(n+3)(n+4)}, \tag{13}$$

which tends towards 6 (instead of the normal value 3) when n tends towards infinity. The distribution of the mean tends rapidly towards normality when n tends towards infinity. Formula (6) cannot be used in this case. Actually it is found that, when n is large,

$$\text{Prob}\{|\,x\,|<|\,t\,|\} = \sqrt{\left(\frac{6}{n\pi}\right)} = 1{\cdot}38n^{-\frac{1}{2}}. \tag{14}$$

It is interesting to compare this probability with the probability which would be found from formula (6), i.e. on the assumption that both estimates were normally distributed. The coefficient of correlation between mean and mid-point is

$$\rho = \sqrt{\left(\frac{6n}{(n+1)(n+2)}\right)} = \sqrt{(6)}\,n^{-\frac{1}{2}} \tag{15}$$

and the pseudo-probability required is approximately

$$\frac{2\sqrt{6}}{\pi}n^{-\frac{1}{2}} = 1{\cdot}56n^{-\frac{1}{2}},$$

which is very little different from the true probability at (14).

In the previous application one of the two estimates (namely the mean) was approximately normally distributed. The mid-point is now compared with another estimate, the variance of which is of the same order of magnitude (in n) as that of the mid-point, but the distribution of which does not tend towards normality with increasing n.

This other estimate is the lowest point in the sample less its mean for $\theta = 0$. This estimate is clearly consistent. Since both estimates are consistent their deviations may be regarded as measured from θ, the unknown centre of the rectangular distribution.

In the range $(-\frac{1}{2}, +\frac{1}{2})$ the distribution of the lowest point is given by $n(\frac{1}{2}-u)^{n-1}du$, the mean value of which is $-k = -\frac{1}{2}(n-1)/(n+1)$. The proposed estimate is $(u+k)$. If the largest value is v the mid-point estimate is $t = \frac{1}{2}(u+v)$. Hence it is required to consider the probability of

$$|u+k| < \tfrac{1}{2}|u+v|. \tag{16}$$

The joint distribution of u and v is given by

$$n(n-1)\,|v-u|^{n-2}\,du\,dv. \tag{17}$$

In the (u, v) plane the inequalities (10) define two 'critical' straight lines given by

$$v-u-2k = 0 \quad \text{and} \quad v+3u+2k = 0.$$

The required probability will be found by integrating (17) over certain areas bounded by these lines and by the lines $u = \pm\frac{1}{2}$, $v = \pm\frac{1}{2}$. The probability is a very complicated function of n which, however, reduces to

$$\tfrac{3}{4}e^{-\frac{2}{3}} - \tfrac{1}{2}e^{-2} = 0{\cdot}317, \tag{18}$$

when $n = \infty$. Accordingly the mid-point is a much closer estimate of θ than the lowest point less its mean for $\theta = 0$.

As regards efficiency, the variances of u and t and the coefficient of correlation between u and t are as follows:

$$\sigma_u^2 = n/\{(n+1)^2(n+2)\}, \quad \sigma_t^2 = 1/\{2(n+1)(n+2)\}, \quad \rho_{ut} = \sqrt{(n+1)}/\sqrt{(2n)},$$

so that, if u and t were normally distributed, the probability that $|u-\bar{u}| < |t|$, from (6), is given by

$$\frac{1}{\pi}\tan^{-1}(2\sqrt{(n+1)}/\sqrt{(n-1)})$$

which, when n tends towards infinity, tends towards 0·352. This is not very different from the true probability 0·317.

A FURTHER ILLUSTRATION: 'CARS IN A TOWN'

At a recent meeting of the Dublin University Mathematical Society, E. Schrödinger suggested the following ingenious problem as an illustration of Pitman's concept of *closeness*. In a town, cars are known to be numbered consecutively from 1. The numbers on r of the cars are noted: the problem is to find the closest estimate of the number of cars in the town. Following is the solution, on Pitman's lines, of Schrödinger's problem.

Let n, the unknown total number, be assumed to be so large* that variation is continuous, i.e. that any car number observed at random has a rectangular frequency distribution. The highest of the r numbers observed, namely w, is a sufficient statistic for n because, when w is known, the remaining $(r-1)$ variates have a joint frequency distribution independent of n; hence all relevant information can be derived from w; the remaining $(r-1)$ observations may be ignored. The cumulative frequency distribution of w is

$$w^r/n^r.$$

By Pitman's theory the closest estimate will be that for which the observed w has the median value, i.e. if $\hat{n}$ is the estimate of n,

$$(w/\hat{n})^r = \tfrac{1}{2} \quad \text{or} \quad \hat{n} = 2^{1/r}w.$$

It is clear that $2^{1/r}w$ has median value n. If $r = 1$, $\hat{n} = 2w$, and if r tends towards ∞, $\hat{n}$ tends towards w, both of which are reasonable. That $\hat{n}$ actually is the closest estimate transpir

* It seems likely that the solution is valid for all values of n.

from a theorem of Pitman (1937) to the effect that if $\hat{n}$ has median value n, and if n' be any other estimating function, then $\hat{n}$ is a closer estimate of n than n', i.e.

$$\text{Prob}\{|\hat{n}-n|<|n'-n|\}>\tfrac{1}{2},$$

provided that a variable z can be found, always of the same sign, so that

$$\hat{n} \quad \text{and} \quad z(n'-\hat{n})$$

are independent.

In the present application suppose that n' is another estimate of n. From considerations of scale, n' must be homogeneous and of degree 1 in the observation w and the other observations $w_1, w_2, \ldots, w_{r-1}$. Hence $n'/\hat{n}$ must be homogeneous and of degree 0 in $w, w_1, w_2, \ldots, w_{r-1}$, and therefore expressible as a function of

$$q_i = w_i/w, \quad i = 1, 2, \ldots, r-1.$$

Now it is obvious that the q_i, and hence $n'/\hat{n}$, are independent of w. Therefore taking z as $1/\hat{n}$, which is always positive, we see that $\hat{n}$ is the closest estimate.

To find an upper limit (in probability) of n, we express the fact that n should not be so great as to render too unlikely the occurrence of the largest number actually observed w. Accordingly, pre-determine a probability α and set

$$(w/n)^r \geqslant \alpha.$$

Hence $$n \leqslant w/\alpha^{1/r}.$$

Example. The number of 30 motor cars are noted and the largest number is found to be 247: to estimate the number of cars in the town and the upper limit of error of the estimate. Here $r = 30$, $w = 247$, so that

$$n = 2^{1/30} \times 247 = 253 \text{ approximately},$$

and $$n \leqslant 247/(0{\cdot}05)^{1/30} = 273 \text{ approximately, if } \alpha = 0{\cdot}05.$$

The latter statement means that the number of cars in the town will be less than 273 unless in taking the particular sample an event, the probability of which was 1/20, occurred.

Summary and conclusion

The criterion of *efficiency* as determined by a comparison of the variances of two estimates of an unknown statistic is identical with the criterion of *closeness* when the joint distribution of the estimates is normal and the criteria will not yield significantly different results in practice when the estimates are estimated from large samples and are consistent. Study of some particular examples suggests that, even when the distribution of the estimates is very different from normal the value of the probability associated with the criterion of closeness may not be very different from what its value would be if normality of the joint distribution were assumed. An application of Pitman's theory of closeness is discussed.

The author is much indebted to Prof. E. S. Pearson for his helpful criticism of this Note, and to Mr N. L. Johnson for pointing out an error in algebra in the first draft submitted.

References

Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics. *Philos. Trans.* A, **122**, 309.

Fisher, R. A. (1935). The Logic of Inductive Inference. *J. Roy. Statist. Soc.* **98**, 39.

Neyman, J. & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika*, **20** A, 175.

Pitman, E. J. G. (1937). The 'closest' estimates of statistical parameters. *Proc. Camb. Phil. Soc.* **33**, 212.

THE RELATION BETWEEN MEASURES OF CORRELATION IN THE UNIVERSE OF SAMPLE PERMUTATIONS

By H. E. DANIELS, *Wool Industries Research Association*

1. Introduction

Recent papers by Hotelling & Pabst (1936), Pitman (1937), Kendall (1938) and Kendall, Kendall & Babington-Smith (1939) discuss the distribution of the correlation coefficient when members of the sample corresponding to the two variates are permuted randomly relative to one another. In the case of rank correlation, the characteristics of the population sampled are generally unknown, and a significance test has to be based on the distribution obtained from the sample in this way.

Hotelling & Pabst prove that as the sample size is increased, Spearman's ρ tends to follow a normal distribution law. Kendall's measure of rank correlation, τ, in which all possible corresponding pairs in two given rankings are assigned marks ± 1 according to whether they agree or differ in order, follows a specially simple distribution law which tends rapidly to the normal form and becomes highly correlated with Spearman's ρ for samples of moderate size.

The present paper discusses the properties of the class of correlation coefficients Γ obtained on replacing Kendall's marks ± 1 by a more general system of scores. By an empirical argument Kendall *et al.* showed it to be likely that the correlation between τ and ρ is $\dfrac{2(n+1)}{\sqrt{[2n(2n+5)]}}$ for all values of the sample size n, and surmised that their joint distribution tends to the bivariate normal form for large n. These results are, in fact, special cases of the relations demonstrated below between two correlation coefficients Γ with different systems of scores.

2. Definition

Consider the two sets of n sample values

$$x_1, x_2, \ldots, x_n, \quad y_1, y_2, \ldots, y_n,$$

both arranged in some given order relative to each other. They may be permuted to give $n!$ different ways of grouping the x's with the y's. Let us assign to each pair (x_i, x_j) what for convenience will be termed a score a_{ij} and to each (y_i, y_j) a score b_{ij}, where

$$a_{ij} = -a_{ji}, \quad b_{ij} = -b_{ji}.$$

Denote by Γ the number
$$\Gamma = \frac{\Sigma a_{ij} b_{ij}}{\sqrt{(\Sigma a_{ij}^2\, \Sigma b_{ij}^2)}},$$

the summation extending over all i and j from 1 to n. Special cases of Γ are Kendall's τ, the product-moment correlation coefficient r and Spearman's ρ, for τ is obtained by definition when $a_{ij}, b_{ij} = \pm 1, j \gtrless i$, r is given when $a_{ij} = x_j - x_i$, $b_{ij} = y_j - y_i$ by virtue of the identity

$$\tfrac{1}{2}\sum_i \sum_j (x_j - x_i)(y_j - y_i) \equiv n \sum_i x_i y_i - \sum_i \sum_j x_i y_j,$$

and ρ is similarly obtained when $a_{ij}, b_{ij} = j - i$.

When the x's are permuted relative to the y's, the scores reappear in a new order with the same or opposite sign and the denominator of Γ remains unaltered, so that in discussing

the distribution of Γ over all permutations it is sufficient to consider the numerator only, which we denote by c.

Write A, B for the matrices of the scores; for example, with $n = 4$,

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & b_{12} & b_{13} & b_{14} \\ -b_{12} & 0 & b_{23} & b_{24} \\ -b_{13} & -b_{23} & 0 & b_{34} \\ -b_{14} & -b_{24} & -b_{34} & 0 \end{bmatrix}.$$

With the x's and y's in the order as written, c is the trace of the matrix product AB' (i.e. the sum of the elements of its leading diagonal), where B' is the transpose of B. The effect of a permutation of the x's, say, is to alter the score matrix of the x pairs to PAP' and the value of c to the trace of $PAP'B'$, where $P = (p_{ij})$ is the appropriate 'permutation matrix' obtained by permuting the columns of the unit matrix. For example, corresponding to the grouping

$$x_4, x_1, x_2, x_3, \quad y_1, y_2, y_3, y_4,$$

the permutation matrix is

$$P = \begin{bmatrix} . & . & . & 1 \\ 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \end{bmatrix}.$$

In terms of the matrix elements c is given by

$$c = \Sigma p_{ji} p_{kl} a_{il} b_{jk},$$

all suffixes being summed from 1 to n.

3. Moments

The distribution of c over all permutations, or the joint distribution of two c's with different systems of scores, is most readily discussed from its moments. The product moment of $c(1)$ and $c(2)$ with scores $a^{(1)}_{ij}$, $b^{(1)}_{ij}$ and $a^{(2)}_{ij}$, $b^{(2)}_{ij}$ respectively is

$$\overline{c(1)\,c(2)} = \frac{1}{n!} S(c(1)\,c(2)),$$

where

$$c(1)\,c(2) = \Sigma p_{ji} p_{kl} p_{sr} p_{ut} a^{(1)}_{il} b^{(1)}_{jk} a^{(2)}_{rt} b^{(2)}_{su},$$

and S denotes summation over all $n!$ possible permutations. Consider the effect of S on each term. The non-vanishing contributions occur when all four p's are 1, and it is first noted that if any of the suffixes i, l, r, t are equal, the corresponding suffixes in j, k, s, u must also be equal for the term not to vanish, since each row of P contains only one non-vanishing element. Terms in which $i = l$ or $r = t$ are of course zero by definition. When, for example, $i = r$ so that r is replaced by i in the expression, we shall call i a *tied suffix*. Other suffixes will be referred to as *free* suffixes.

As regards their contribution to S the terms may be classified according to the number of tied suffixes in i, l, r, t as follows.

(i) *No tied suffixes.* When the four p's are each unity, four rows and columns of P are assigned and there are $(n-4)!$ ways of filling the remaining positions. Such terms therefore contribute

$$(n-4)!\,\Sigma' a^{(1)}_{il} a^{(2)}_{rt}\,\Sigma' b^{(1)}_{jk} b^{(2)}_{su}$$

to the total sum, where Σ' denotes summation over all values of the suffixes which are not

equal. Let us consider some properties of $\Sigma' a_{il} a_{rt}$ and $\Sigma a_{il} a_{rt}$ and similar expressions with tied suffixes, the second expression being summed over *all* values of the suffixes. (Superscripts (1) and (2) are understood throughout.)

(1) $\Sigma a_{ii} a_{rt} = 0$, $\Sigma' a_{ii} a_{rt} = 0$, and so on.

(2) $\Sigma a_{il} a_{il} = \Sigma' a_{il} a_{il}$.

(3) $\Sigma a_{il} a_{it} = \Sigma' a_{il} a_{it} + \Sigma' a_{il} a_{il}$.

(4) $\Sigma a_{il} a_{rt} = 0$, $\Sigma' a_{il} a_{rt} = 0$. The first is true because $\Sigma a_{il} = 0$, and the second follows from the fact that

$$\Sigma a_{il} a_{rt} = \Sigma' a_{il} a_{rt} + \Sigma' a_{il} a_{it} + \Sigma' a_{il} a_{ri} + \Sigma' a_{il} a_{il} + \Sigma' a_{il} a_{rl} + \Sigma' a_{il} a_{il} + \Sigma' a_{il} a_{li},$$

the terms on the right after the first cancelling in pairs.

The contribution to S of terms with no tied suffixes is therefore zero.

(ii) *One tied suffix.* For the term not to vanish it is necessary to assign three rows and columns of P, and the contribution to S from such terms is

$$4(n-3)!\,\Sigma' a_{il}^{(1)} a_{it}^{(2)} \Sigma' b_{jk}^{(1)} b_{ju}^{(2)},$$

the factor 4 arising from the fact that the same contribution is obtained by tying the suffixes in the four possible ways.

(iii) *Two tied suffixes.* The contribution to S is similarly found to be

$$2(n-2)!\,\Sigma' a_{il}^{(1)} a_{il}^{(2)} \Sigma' b_{jk}^{(1)} b_{jk}^{(2)}.$$

Terms containing more than two tied suffixes give zero contributions to S, and finally, substituting for Σ' the appropriate Σ expressions, we find

$$\overline{c(1)\,c(2)} = \frac{4}{n(n-1)(n-2)} (\Sigma a_{il}^{(1)} a_{it}^{(2)} - \Sigma a_{il}^{(1)} a_{il}^{(2)}) (\Sigma b_{jk}^{(1)} b_{ju}^{(2)} - \Sigma b_{jk}^{(1)} b_{jk}^{(2)}) + \frac{2}{n(n-1)} \Sigma a_{il}^{(1)} a_{il}^{(2)} \Sigma b_{jk}^{(1)} b_{jk}^{(2)}.$$

The moments of higher order can be obtained by a similar procedure, but the expressions rapidly become unwieldy.

4. The correlation between Kendall's τ and Spearman's ρ

As a first application of the formula we consider the correlation between τ and ρ over all permutations of the sample values. The scores for τ and ρ respectively are

$$a_{ij}^{(1)},\ b_{ij}^{(1)} = \pm 1,\ 0 \quad \text{when} \quad j \gtrless i,\ j = i,$$

$$a_{ij}^{(2)},\ b_{ij}^{(2)} = j - i.$$

The following results are easily derived

$$\sum_{l=1}^{n} a_{il}^{(1)} = n+1-2i, \quad \sum_{l=1}^{n} a_{il}^{(2)} = \tfrac{1}{2}n(n+1-2i),$$

$$\sum_{i=1}^{n} \sum_{l=1}^{n} \sum_{t=1}^{n} a_{il}^{(1)} a_{it}^{(2)} = \frac{n^2(n^2-1)}{6}, \quad \sum_{i=1}^{n} \sum_{l=1}^{n} a_{il}^{(1)} a_{il}^{(2)} = \frac{n(n^2-1)}{3},$$

and the same results hold for the b's. Substitution in the formula then gives

$$\overline{c(1)\,c(2)} = \frac{n^2(n-1)(n+1)^2}{9}.$$

Again, $$\sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{t=1}^{n} a_{il}^{(1)}a_{it}^{(1)} = \frac{n(n^2-1)}{3}, \quad \sum_{i=1}^{n}\sum_{l=1}^{n} a_{il}^{(1)}a_{il}^{(1)} = n(n-1)$$

and $$\sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{t=1}^{n} a_{il}^{(2)}a_{it}^{(2)} = \frac{n^3(n^2-1)}{12}, \quad \sum_{i=1}^{n}\sum_{l=1}^{n} a_{il}^{(2)}a_{il}^{(2)} = \frac{n(n^2-1)}{3},$$

from which it is found that

$$\overline{c(1)^2} = \frac{2n(n-1)(2n+5)}{9}, \quad \overline{c(2)^2} = \frac{n^4(n-1)(n+1)^2}{36}.$$

The required correlation is therefore

$$R_{\rho\tau} = \frac{2(n+1)}{\sqrt{[2n(2n+5)]}}$$

which is the result anticipated by Kendall *et al.* It should be noted that they use the quantities $\Sigma = \frac{1}{2}c(1)$ and

$$S(d^2) = \frac{n(n^2-1)}{6} - \frac{c(2)}{n}$$

in place of $c(1)$ and $c(2)$.

5. Transformation of the sample values

If the scales of the x's and y's are distorted by a transformation and the product moment correlation coefficient r is recalculated on the transformed sample, a new value r' is obtained. In particular, the x's and y's may be readjusted to be at equidistant intervals, and then the new value of r is Spearman's ρ. The formula for $\overline{c(1)\,c(2)}$ can be used to find the correlation over all sample permutations between the values of r on the same sample before and after such a transformation. Distinguishing by primes the sample values after transformation, the scores are

$$a_{ij}^{(1)} = x_j - x_i, \quad b_{ij}^{(1)} = y_j - y_i,$$
$$a_{ij}^{(2)} = x'_j - x'_i, \quad b_{ij}^{(2)} = y'_j - y'_i.$$

Then $$\sum_{l=1}^{n} a_{il}^{(1)} = n(\bar{x} - x_i), \quad \sum_{i=1}^{n}\sum_{l=1}^{n}\sum_{t=1}^{n} a_{il}^{(1)}a_{it}^{(2)} = n^2 \sum_{i=1}^{n}(x_i-\bar{x})(x'_i-\bar{x}'),$$

$$\sum_{i=1}^{n}\sum_{l=1}^{n} a_{il}^{(1)}a_{il}^{(2)} = \sum_{i=1}^{n}\sum_{l=1}^{n}(x_l-x_i)(x'_l-x'_i) = 2n\Sigma(x_i-\bar{x})(x'_i-\bar{x}')$$

by the identity previously quoted. Using these and similar formulae we find

$$\overline{c(1)\,c(2)} = \frac{4n^2}{n-1}\Sigma(x-\bar{x})(x'-\bar{x}')\Sigma(y-\bar{y})(y'-\bar{y}'),$$

$$\overline{c(1)^2} = \frac{4n^2}{n-1}\Sigma(x-\bar{x})^2\Sigma(y-\bar{y})^2; \quad \overline{c(2)^2} = \frac{4n^2}{n-1}\Sigma(x'-\bar{x}')^2\Sigma(y'-\bar{y}')^2,$$

and hence the correlation between r and r' is

$$R_{rr'} = r_{xx'}r_{yy'},$$

where $r_{xx'}$ and $r_{yy'}$ are the correlation coefficients between old and new values of x and y respectively.

6. Tendency to normal form for large n

It will now be shown for a large class of score systems a_{ij} that c, and hence Γ, tends with increasing n to be normally distributed, and moreover, that the joint distribution of any pair of such Γ's tends to the bivariate normal form.

The pth order product moments of the joint distribution of $c(1)$ and $c(2)$ are sums of terms containing

$$\Sigma' a_{gh} a_{ij} a_{kl} \ldots \Sigma' b_{rs} b_{tu} b_{vw} \ldots$$

or similar expressions in which arbitrary groups of suffixes within the Σ''s are tied, each Σ' involving products of p scores which may belong either to systems (1) or (2). Every such Σ' is in turn a linear combination of the corresponding Σ having the same suffixes and other Σ's in which additional tied suffixes are introduced. No Σ may contain a pair of free suffixes attached to one score, for it would then vanish by virtue of the fact that $\Sigma a_{ij} = 0$.

The even order product moments are first discussed. Let $p = 2m$. Consider a Σ in which the $2m$ scores are divided into m pairs each having one tied suffix, so that there are in all $3m$ independent suffixes, e.g.

$$\Sigma a_{ij} a_{ik} a_{lr} a_{ls} a_{tu} a_{tv} \ldots.$$

It may be written as

$$(\Sigma a_{ij}^{(1)} a_{ik}^{(1)})^{\lambda} (\Sigma a_{ij}^{(1)} a_{ik}^{(2)})^{\mu} (\Sigma a_{ij}^{(2)} a_{ik}^{(2)})^{\nu},$$

where $\lambda+\mu+\nu = m$ and λ, μ, ν are the number of times the scores are paired in the combinations indicated.

As is always possible, suppose the numerically largest value of a_{ij} to be made equal to unity. We now impose the condition that $\Sigma a_{ij} a_{ik}$ is of order n^3 whether a_{ij} and a_{ik} belong to the same or different systems of scores. This is satisfied, when $\max a_{ij} = 1$, by τ and ρ, and also by r provided the sample is not an unusual one. With this condition, it is seen that Σ's of the above type are of order n^{3m}.

It is next observed that all other ways of tying suffixes give Σ's of lower order of magnitude. For the order of magnitude of the bracket is not reduced on replacing each a_{ij} by $+1$; consequently if further suffixes are tied the order of Σ is made less than n^{3m} since there are fewer than $3m$ summations from 1 to n. It follows that the dominant term in a Σ' is the corresponding Σ having the same array of suffixes.

Moreover, every non-vanishing Σ involving $3m$ independent suffixes can only be a permutation of the type illustrated, while those with more than $3m$ different suffixes must all vanish. This is made clear by considering how the $3m$ suffixes can be arrayed between the $2m$ scores. Begin by assigning $3m$ different suffixes at random among the $4m$ available places. At least m scores will receive their full complement of suffixes all which will be different. There cannot be more than m such completed scores, for if Σ is not to vanish, at least one suffix of each complete pair must be tied and this can only be done by repeating one suffix from every complete pair in each of the remaining places to be filled, of which there are only m. We are thus led to a permutation of the type of Σ discussed above. If there had been more than $3m$ different suffixes to begin with, there would not have remained sufficient empty places to prevent the existence of at least one score with a pair of free suffixes, and so all Σ's with more than $3m$ different suffixes must vanish.

Any $2m$th product moment is the sum of terms like

$$\frac{(n-f)!}{n!} A \Sigma' a_{ij} a_{kl} \ldots \Sigma' b_{rs} b_{tu} \ldots,$$

where f is the number of independent suffixes in the Σ''s and A is a coefficient which is of

unit order as far as n is concerned. From the preceding argument, the maximum value of f is $3m$, in which case the term is of order $n^{-3m} \times n^{3m} \times n^{3m} = n^{3m}$. When $f \leqslant 3m-1$ the order of the term is not greater than $n^{-3m+1} \times n^{3m-1} \times n^{3m-1} = n^{3m-1}$ and such terms may therefore be neglected. Write

$$h_{11} = \Sigma a_{ij}^{(1)} a_{ik}^{(1)} \Sigma b_{lu}^{(1)} b_{lv}^{(1)}, \quad h_{12} = \Sigma a_{ij}^{(1)} a_{ik}^{(2)} \Sigma b_{lu}^{(1)} b_{lv}^{(2)}, \quad h_{22} = \Sigma a_{ij}^{(2)} a_{ik}^{(2)} \Sigma b_{lu}^{(2)} b_{lv}^{(2)}.$$

Then if terms of lower order of magnitude are neglected, the even product moment

$$\mu_{r,s} = \overline{c(1)^r c(2)^s}, \quad r+s = 2m$$

is given by the sum of terms like

$$n^{-3m} A_{\lambda,\mu,\nu} h_{11}^{\lambda} h_{12}^{\mu} h_{22}^{\nu}; \quad 2\lambda + \mu = r, \quad \mu + 2\nu = s,$$

over all possible values of λ, μ, ν. The coefficient $A_{\lambda,\mu,\nu}$, which is the number of ways in which $h_{11}^{\lambda} h_{12}^{\mu} h_{22}^{\nu}$ can arise, is calculated as follows. Consider a Σ whose array of suffixes is such that it can be factorized as $(\Sigma a_{ij}^{(1)} a_{ik}^{(1)})^{\lambda} (\Sigma a_{ij}^{(1)} a_{ik}^{(2)})^{\mu} (\Sigma a_{ij}^{(2)} a_{ik}^{(2)})^{\nu}$. Its suffix pairs can be permuted within the sets of scores (1) and (2) in $r!\,s!$ ways, but of these $\lambda!(2!)^{\lambda}\,\mu!\,\nu!(2!)^{\nu}$ give essentially the same Σ. The suffixes within pairs attached to each score may also be rearranged in 2^{2m} ways without affecting the result, and so

$$A_{\lambda,\mu,\nu} = \frac{r!\,s!\,2^{2m}}{\lambda!\,\mu!\,\nu!\,2^{\lambda+\nu}} = \frac{r!\,s!\,2^{m+\mu}}{\lambda!\,\mu!\,\nu!}.$$

The calculation of the even order product moment $\mu_{r,s}$ for large n is in fact tantamount to selecting the coefficient of $t^r t^s / r!\,s!$ in

$$\frac{2^m}{n^{3m} m!} (h_{11} t_1^2 + 2h_{12} t_1 t_2 + h_{22} t_2^2)^m.$$

Finally, we dispose of the odd moments. In certain cases, such as for example the joint distribution of τ and ρ, they all vanish by symmetry. But even in the general case it can be shown that the odd moments are negligible to the order of magnitude $n^{-\frac{1}{2}}$.

A Σ containing $2n+1$ scores cannot have more than $3m+1$ different suffixes. For if there were $3m+2$, let them first be assigned to the $4m+2$ available places; at least $m+1$ scores will receive complete pairs of suffixes, and the remaining m empty places cannot be filled in any way which avoids one score having a free pair of suffixes. Hence as before the order of magnitude of any $(2m+1)$th moment is at most $n^{-(3m+1)} \times n^{3m+1} \times n^{3m+1} = n^{3m+1}$

The $2m$th moments were shown to be of order n^{3m}, consequently if we define

$$\gamma(1) = n^{-\frac{3}{2}} c(1), \quad \gamma(2) = n^{-\frac{3}{2}} c(2),$$

the joint distribution of $\gamma(1)$ and $\gamma(2)$ has all its even moments of unit order, and by the result just proved all its odd moments are of order $n^{-\frac{1}{2}}$ and may therefore be neglected to that order. Reverting to $c(1)$, $c(2)$, it is seen that the moment-generating function of their joint distribution tends in the limit to the form

$$\exp \frac{2}{n^3} (h_{11} t_1^2 + 2h_{12} t_1 t_2 + h_{22} t_2^2).$$

Hence $c(1)$ and $c(2)$ tend to be normally distributed with variances $\frac{4}{n^3} h_{11}$, $\frac{4}{n^3} h_{22}$ and correlation

$$\frac{h_{12}}{\sqrt{(h_{11} h_{22})}} = \frac{\Sigma a_{ij}^{(1)} a_{ik}^{(2)}}{\sqrt{(\Sigma a_{ij}^{(1)} a_{ik}^{(1)} \Sigma a_{ij}^{(2)} a_{ik}^{(2)})}} \frac{\Sigma b_{lu}^{(1)} b_{lv}^{(2)}}{\sqrt{(\Sigma b_{lu}^{(1)} b_{lv}^{(1)} \Sigma b_{lu}^{(2)} b_{lv}^{(2)})}}.$$

The Γ's similarly tend to a bivariate normal distribution with the same correlation, but with variances,

$$\frac{4}{n^3}\frac{\Sigma a_{ij}^{(1)}a_{ik}^{(1)}\Sigma b_{lu}^{(1)}b_{lv}^{(1)}}{\Sigma a_{ij}^{(1)}a_{ij}^{(1)}\Sigma b_{lu}^{(1)}b_{lu}^{(1)}},\quad \frac{4}{n^3}\frac{\Sigma a_{ij}^{(2)}a_{ik}^{(2)}\Sigma b_{lu}^{(2)}b_{lv}^{(2)}}{\Sigma a_{ij}^{(2)}a_{ij}^{(2)}\Sigma b_{lu}^{(2)}b_{lu}^{(2)}}.$$

Our proof rests on the assumption that $\Sigma a_{ij}a_{ik}$ and $\Sigma b_{lu}b_{lv}$ are of order n^3, where the individual a_{ij}'s and b_{lu}'s may belong to either score system. But if that is true, it follows that expressions like Σa_{ij}^2 must be of order n^2, for they cannot be made to exceed that order on replacing a_{ij} by $+1$, and their order cannot be less than n^2 since

$$\Sigma a_{ij}^2 - \frac{1}{n}\Sigma a_{ij}a_{ik} = \Sigma(a_{ij}-\bar{a}_i)^2 \geqslant 0,$$

where $\bar{a}_i = \frac{1}{n}\sum_{j=1}^{n} a_{ij}$. Consequently the variances of the Γ's decrease like n^{-1}. The correlation between the Γ's tends, however, to a value independent of n in the limit.

SUMMARY

The properties of a general class of correlation coefficients Γ, which includes the product-moment correlation coefficient r, Spearman's ρ and Kendall's τ, are discussed. A direct proof is given of the formula tentatively suggested by Kendall for the correlation between ρ and τ when the sample is permuted in all possible ways. The effect of a transformation of the sample values is also considered. It is shown that under certain general conditions, the joint distribution of two different Γ's, calculated on all possible permutations of the sample values, tends with increasing sample size to the bivariate normal form with variances inversely proportional to the sample size and correlation independent of it.

REFERENCES

HOTELLING, H. & PABST, M. R. (1936). Rank correlation and tests of significance involving no assumptions of normality. *Ann. Math. Statist.* 7, 29.

PITMAN, E. J. G. (1937). Significance tests which may be applied to samples from any populations. II. The correlation coefficient test. *J.R. Statist. Soc. Suppl.* **4**, 225.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81.

KENDALL, M. G., KENDALL, S. F. H. & BABINGTON-SMITH, B. (1939). The distribution of Spearman's coefficient of rank correlation in a universe in which all rankings occur an equal number of times. *Biometrika*, **30**, 251.

THE GROWTH, SURVIVAL, WANDERING AND VARIATION OF THE LONG-TAILED FIELD MOUSE, *APODEMUS SYLVATICUS*

By H. P. HACKER and H. S. PEARSON

I. GROWTH. By HELGA S. PEARSON

CONTENTS

1. Introduction

Pioneer work in trapping field mice alive has been done in this country by Charles Elton, his colleagues and Dennis Chitty. Reference to similar work in other countries, especially in the United States, will be found in their papers and in the useful lists of recent literature in the *Annual Reports of the Bureau of Animal Population, Oxford University*. We record with gratitude the benefit we have derived from their labours and devices.

Our own interest lay primarily in comparing, by simple statistical methods, populations from species of mammals distributed more or less continuously over great distances. These comparisons were to be based on skeletal measurements and proportions. We were limited to our own country and to a ubiquitous mammal easily caught in large numbers; thus field mice seemed to be the only choice, although for skeletal measurements these small rodents present great difficulties. Our Oxford advisers recommended the long-tailed field or wood mouse, *Apodemus sylvaticus sylvaticus*, as more or less ubiquitous and easy to catch. We decided to follow in their footsteps, and start by trapping these mice alive, marking them, setting them free and trapping again, in the hope of getting some first-hand experience of how far they wander, and whether they are confined to particular types of country or can be regarded as a single population continuous throughout England.

Even if a whole species can be regarded as a continuous population it is clear that there can be no such thing as 'random mating' within it, if the individuals rarely wander more than a few hundred yards. There must be regional inbreeding, though there may be no boundaries between the regions. Can any regional differences be detected between field mice if measurable characters are compared by statistical methods?

The trapping of *Apodemus* alive is a fascinating pursuit. Many more questions arise from it than can be answered perhaps in a lifetime, and it has proved hard to hold firmly to the pursuit of our original problem. In trapping over a number of seasons it has been easy to collect a vast quantity of facts, and these, when arranged and digested, cry out for the

collection of a vast quantity more. It has seemed advisable, however, to try to publish some of them at this stage, in the hope that they may prove of interest or use to other workers, if we ourselves are not able to follow them up much further.

The data may be arranged, though with much overlapping, under the four headings: Growth, Variation, Wandering, Survival. We hope to publish some selected facts appropriate to these headings in a short series of papers of which this is the first.

2. TRAPPING TECHNIQUE

That many of our present records are not suitable for statistical treatment is due to our ignorance when first trapping of the best method of laying out our traps. Only now, after five seasons' trapping following on the work of Elton and of Chitty, are we beginning to understand the conditions under which large numbers of *Apodemus* can best be caught.

There are four main limiting factors:

(i) *Man hours.* Elton *et al.* (1931, p. 661) stressed the number of 'man hours' needed for setting, visiting, and resetting large numbers of traps, and for dealing with the mice that are found in them. The largest number of traps we ourselves have ever set out at a time was 96, and 48 mice was our largest night's catch, of which 4 were bank voles and 44 *Apodemus*. Our method was to pick up all the occupied traps, replace them with spare traps, and take them home. There we identified or marked the mice, weighed and measured them, placed them in cages, and wrote up our day's records. Then we rebaited the traps and renewed the food in the attached nest boxes (see p. 145), so that they were ready for taking out to replace others next day. Our biggest catches meant a long, intensely hard day's work for three practised people; had there been more of us we could have put out more traps covering a wider area and caught more mice, and so have had more adequate data for statistical treatment.

(ii) *Weather.* Setting out a large number of traps does not necessarily mean catching a large number of mice. On some nights we have had all our traps out in likely places without a single catch. Weather conditions undoubtedly play an all important part in keeping mice from wandering freely, and this in itself needs a carefully planned investigation. We have learnt from experience that a night of blustering south-west wind is a good trapping night and that, if snow is on the ground, it is of little use leaving out the traps when once the mice living in their immediate neighbourhood have been caught; but we are not sure of the relative effect on the numbers caught of wind, moon, temperature, rain or hoar frost.

(iii) *Arrangement of traps.* Only a certain number of mice will be within reach of any one trap under any particular set of weather conditions, and a great number of traps in a dense 'scatter' will catch no more mice under those conditions than few traps more widely spread. A line of traps will drain a wider area than an equal number arranged in a grid, while a hollow square may catch all the mice within it just as easily as a grid covering the same area. With experience some idea may be gained of a good arrangement, but this must still depend to a large extent on a further limiting factor:

(iv) *Habitat.* Not all places are equally frequented by *Apodemus*, yet this cannot easily be attributed to obvious ecological differences. We have caught mice in large numbers in very varied habitats, whereas a comparatively uniform plant community may have good and bad patches. We are inclined to suspect that any kind of overhead cover, such as dense bracken, heather or thorny scrub, is of as great importance as food or soil factors, but much more work is needed to discover where the mice live as apart from where they wander, and

whether or no they live in large communities in the winter and separate before the breeding season. Light might be thrown on food preferences by analyses of stomach contents and by experimental feeding.

The little information we have on all these matters, together with further notes on our methods of trapping and marking, we hope to publish in our papers on survival and wandering.

3. Measurements

For our future comparison of populations from different localities we needed the skeletons of fully grown mice. In order to find out at what time of year a mouse population contains the largest proportion of such mice, we studied the growth changes of our living population, weighing and measuring each mouse, and recording the numbers of mice of each size in each month; we were also able to keep individual growth records of marked mice. The data thus collected are analysed in the last three sections of this paper.

We had often to handle a large number of live mice in a very little time, so we limited ourselves in their case to two comparatively easy measurements: (i) the weight, (ii) the length of the right hind foot. For any mouse found dead in a trap, or that died in captivity, we made further records of: (iii) the total length of head and trunk, (iv) the length of tail, (v) the weight of stomach, (vi) the weight of the reproductive system in the male, or of visible embryos in any pregnant female.

We have made these same measurements on all mice trapped for skeletons in other localities, and have so gained a first rough index to regional differences. So far we have carried out this regional trapping near Westerham in Kent, Hampden in Bucks, Swanage in Dorset, and St Mawes in Cornwall. We trapped in each of the years 1937, 1938 and 1939, at the very end of March or beginning of April, a time when winter males are past their period of maximum weight increase (p. 155), and when young mice are in most years soon likely to appear in the traps. These dead mice together form the adult population from which the data of Table 1 (p. 139) were calculated. In Fig. 1 their measurements are added to those on dead mice caught at Holwood Park and Downe in Kent, and at Cobbler's Hill in Buckinghamshire, at other times of the year. In a future paper we hope to give an analysis of the local variation shown by this material, together with such further information on growth as can be gained from correlating the external and skeletal measurements of the series of dead mice of all ages from Holwood.

To facilitate statistical analysis the records of each mouse, dead or alive, were entered straightway on to separate index cards, the reference number on each card being that of one individual mouse.

Live mice. It was easy to weigh the live mice. We let them drop from the trap into a large net bag, and from there transferred them by hand to a small, narrow bag in which they could be tied firmly and placed on the balance. The feet were not quite so easy to measure accurately, but we coaxed the mice by gentle pressure from the small bag into a cupped hand, one of us holding them there with the right-hind leg projecting between forefinger and thumb, while another (always the present writer) stretched the hind foot at right angles over the holder's forefinger and measured it from heel to fleshy tip of central digit with a pair of screw callipers with Vernier scale. In this technique we attained a degree of accuracy represented by a standard error for a single measurement of 0·18 mm.*

* This is the S.D. for repeat measurements on the same mouse by one observer, based on 10 or more observations on each of 50 adult male mice, the average foot-length being 23·0 mm.

We sexed each mouse while holding it, and noted whether a female's vulva was perforate. We also placed it in a glass inspection jar—a barrel-shaped lamp glass with one opening sealed and the other wide enough to admit a hand—and noted the amount of yellow streak on the hairs of the chest, a very variable feature.

Post-mortem errors. That favourite measurement of systematists, the combined length of head and trunk, is subject to very great observational error. Indeed, a wide range of error must be allowed for in all external measurements on dead mice, taken as they usually are at varying intervals after death.* Loss of body weight due to drying or decomposition takes place slowly, and we found no appreciable change in 24 hr., but the feet tend to dry and shrink before this. Length of foot, and length of head and trunk, will both be shorter if measured during the period of *rigor mortis* than either before or after; they vary with the stage of rigor and the amount of massage given before measuring, though in extreme rigor it is doubtful if the massage has much effect. As far as dead mice are concerned we hope to obtain more accurate information about growth from skeletal measurements, since these are of a more stable nature and less liable to observational error.

Table 1. *Above: The variation in five measurements on a number* (N.) *of dead adult or nearly adult* Apodemus *caught in four localities between* 28 *March and* 13 *April* 1937, 1938 *and* 1939. *Mean* (M.), *standard deviation* (S.D.), *and coefficient of variation* (C.V.). *Below: The correlation between these measurements; correlation coefficient* (C.C.).

Measurement	Male				Female			
	N.	M.	S.D.	C.V.	N.	M.	S.D.	C.V.
Weight (g.)	161	23·4	2·38	10·2	115	18·8	2·20	11·7
Head and trunk length (mm.)	162	88·6	4·84	5·5	121	84·1	4·89	5·8
Tail length (mm.)	142	90·3	5·27	5·8	107	86·3	4·96	5·7
Right hind foot length (mm.)	162	22·9	0·82	3·6	121	22·3	0·69	3·1
Weight of reproductive organs (g.)	149	2·0	0·41	20·5				

Correlation	Male		Female	
	N.	C.C.	N.	C.C.
Head and trunk length with weight	161	0·59	115	0·53
Head and trunk length with tail length	142	0·51	107	0·46
Head and trunk length with hind foot length	162	0·56	121	0·47
Tail length with hind foot length	142	0·57	107	0·38
Weight with weight of reproductive organs	149	0·66		

Weight as a measure of growth. The easiest measurement of a live mouse, its weight, is unfortunately not a very satisfactory measure of growth. Apart from seasonal changes, it alters considerably from one day, or even hour, to another. This must be partly due to the condition of the stomach and bladder, but we suspect that a rapid fall in actual body weight

* Sumner (1927) has shown how great this error may be, by comparing measurements made on the American deer-mouse *Peromyscus*, (*a*) by different observers, (*b*) by the same observer at different times.

can also occur after a period of prolonged activity. Frequent trapping, and trapping during the breeding season, may also interfere with weight increase.

Some idea of the extent to which the weight of a mouse is correlated with its length, and therefore indirectly with its growth, is given by the scatter diagram of Fig. 1. We were not able to measure the length of our live mice, but this diagram is based on the measurements of the dead male mice referred to on p. 138. It must be examined with reserve, however, as the number of young is few, and the mice were caught at varying times

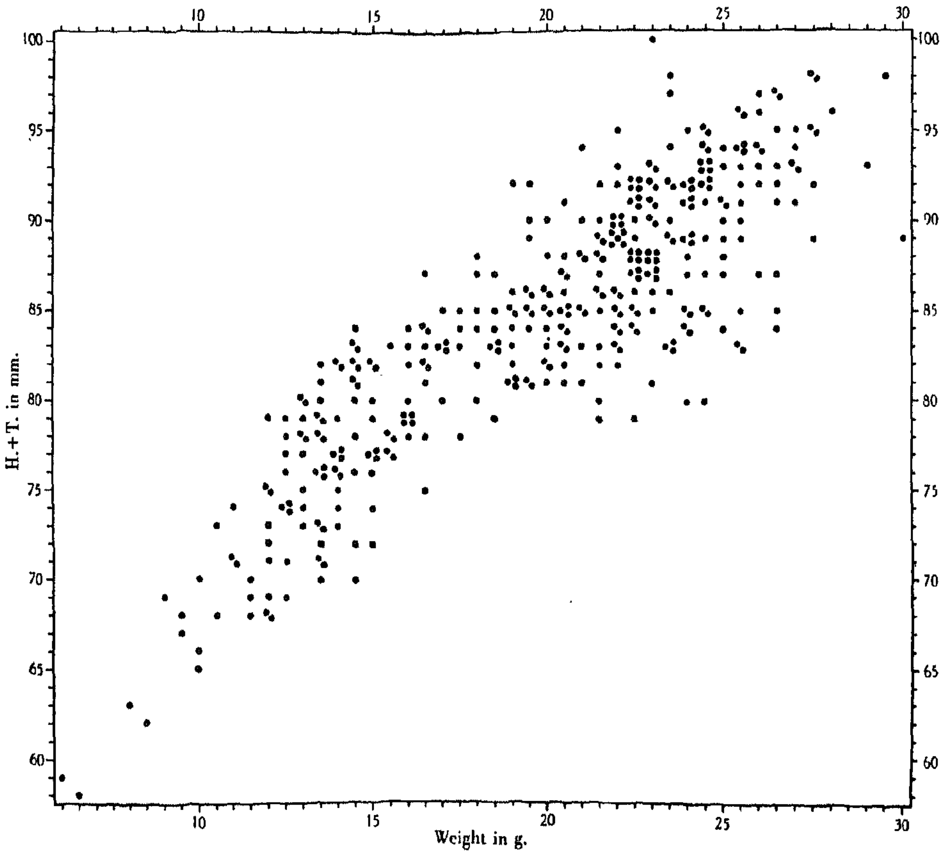


Fig. 1. Correlation between weight and length of head and trunk in 367 dead male *Apodemus* caught in various localities at various times of year.

of year in several widely separated localities, the young in different localities from most of the adults; and, as we hope to show in this and a later paper, weight varies with season and size with locality independently of the age of a mouse. In the mixed batch of *adult* or *nearly adult* mice from Westerham, Hampden, Swanage and St Mawes (included in the upper part of the diagram), we have found the coefficient of correlation between these two measurements to be $r = 0{\cdot}59$ for 161 males and $r = 0{\cdot}53$ for 115 females (Table 1). To calculate comparable coefficients for growing mice would call for far larger numbers in the lower weight groups, and for this reason, and because of the mixed nature of the material,

we have thought it inadvisable to fit a curve to the diagram.* Any such curve would tend to flatten as growth ceases, increase of length with increase in weight becoming solely due to the size variability of the fully grown population. For apart from all the sources of variation mentioned, weight and length must vary not only with age but also with individual character. We have been unable to separate these two sources in growing mice, and without more frequent individual records (and these would probably interfere with growth) it does not seem possible to be certain that any one live mouse is fully grown. Records taken throughout the year suggest that, in the single locality of Holwood Park, 20·0 g. in a male and 17·0 g. in a female can be taken as arbitrary but useful minima below which most of the mice may be expected to be still growing, except in midwinter when heavier adults often fall to lower levels than this (p. 154). It will be seen from Fig. 1 that males of over 20·0 g. range from under 80 to 100 mm. in length, while those of less than 20·0 g. can range as high as 90 mm.

A measure of the variability of weight in *adult* mice from scattered localities is given by the coefficients of variation (100 × standard deviation/mean) which we obtained from the mixed batch mentioned above. These coefficients are given in Table 1, where they are seen to be nearly double the coefficients for length of head and trunk, although the variability in this case includes much greater observational errors. Ruger (1933) found coefficients for weight as high as 15·93 and 14·77 in his human material (corrected to the standard ages of 40·5 years for men and 32·5 years for women), while his coefficients for stature, sitting height, and span are less than ours for head and trunk length, and for tail length.

Summary of factors influencing weight.

(1) Age.
(2) Inherent† size of mouse.
(3) Locality.
(4) Season: (i) fat deposit; (ii) condition of reproductive organs and, in females, presence of embryos.
(5) Prolonged activity.
(6) Trapping interference.
(7) Content of stomach and intestines.
(8) Content of bladder.
(9) Disease (see footnote to p. 158).

Tail length. The length of the tail from anus to tip, excluding the terminal pencil of hairs, is another conventional record of systematists. It is more easily taken on a dead animal than is the length of the head and trunk, and is less subject to observational errors, but as the tail is liable to shortening through injury, our number of records is not so great. The average length of tail in adult mice is very similar to the average length of head and trunk, and this is true for mice of all ages found in the traps, indicating that the tail continues to lengthen as long as the body is lengthening. On the other hand, each varies

* A clear analysis of the difficulties of fitting curves to comparable but much more numerous human measurements has been given by Ruger & Stoessiger (1927) and Ruger (1933). From measurements of fifteen characters taken in 1884 by Francis Galton on over 7000 men, women and children, they computed means, standard deviations, correlation coefficients, correlation ratios and regression equations, and constructed graphs of the regression lines. These papers, and the analytical methods of the biometric school in general, are well worth the consideration of those biologists who attempt to base systematic classification, and even evolutionary theory on so-called *allometric growth curves*.

† *Inherent*. Def. 2 in *Shorter O.E.D.*: 'Existing in something as a permanent attribute or quality'.

to a considerable extent independently of the other (for adult mice, $r = 0{\cdot}51$ in males and $r = 0{\cdot}46$ in females), so that neither can be taken as an index to the other in any particular mouse. Fig. 2 is a correlation diagram for the two characters derived from the same series of dead male mice that formed the basis of Fig. 1. The same qualifications, due to heterogeneity of the data, apply as were referred to on p. 140.

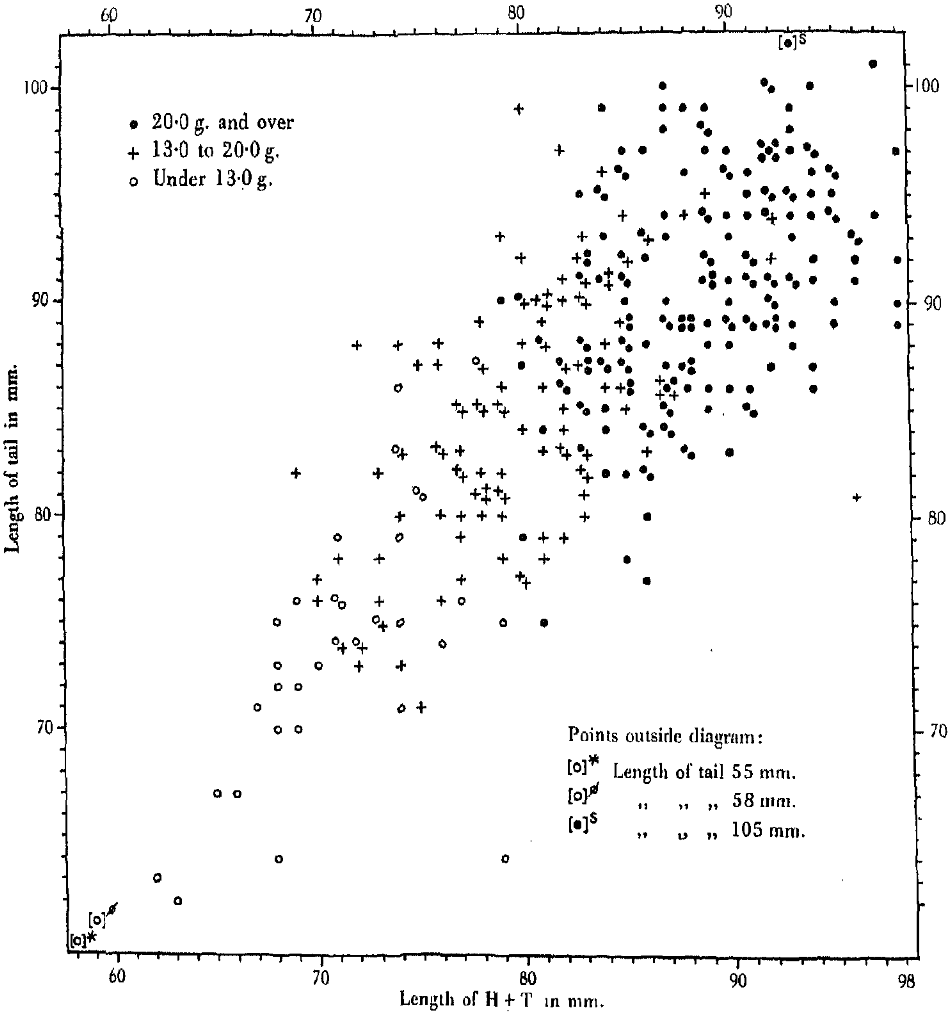


Fig. 2. Correlation between length of tail and length of head and trunk in 322 dead male *Apodemus* caught in various localities at various times of year.

An exceptionally short tail in any weight group may indicate an early loss of tip not perceptible or not noticed on measuring. If *Apodemus* is caught by the end of the tail, the skin easily slips off, if the grip on the bone is not tight, often allowing the mouse to escape; the protruding bone soon dries up and breaks off. If the loss is of any length—and more than half the tail may be lost in this way—it is easily perceptible, as the tail ends bluntly without the terminal pencil of hairs, and sometimes with a short length of exposed bone; in such cases no measurement was made.

Length of hind foot. From our measurements on live *Apodemus* we have found that the hind foot reaches maximum length long before the mouse stops increasing in weight, so that this length, within the error of measurement (see p. 138), early becomes a fixed

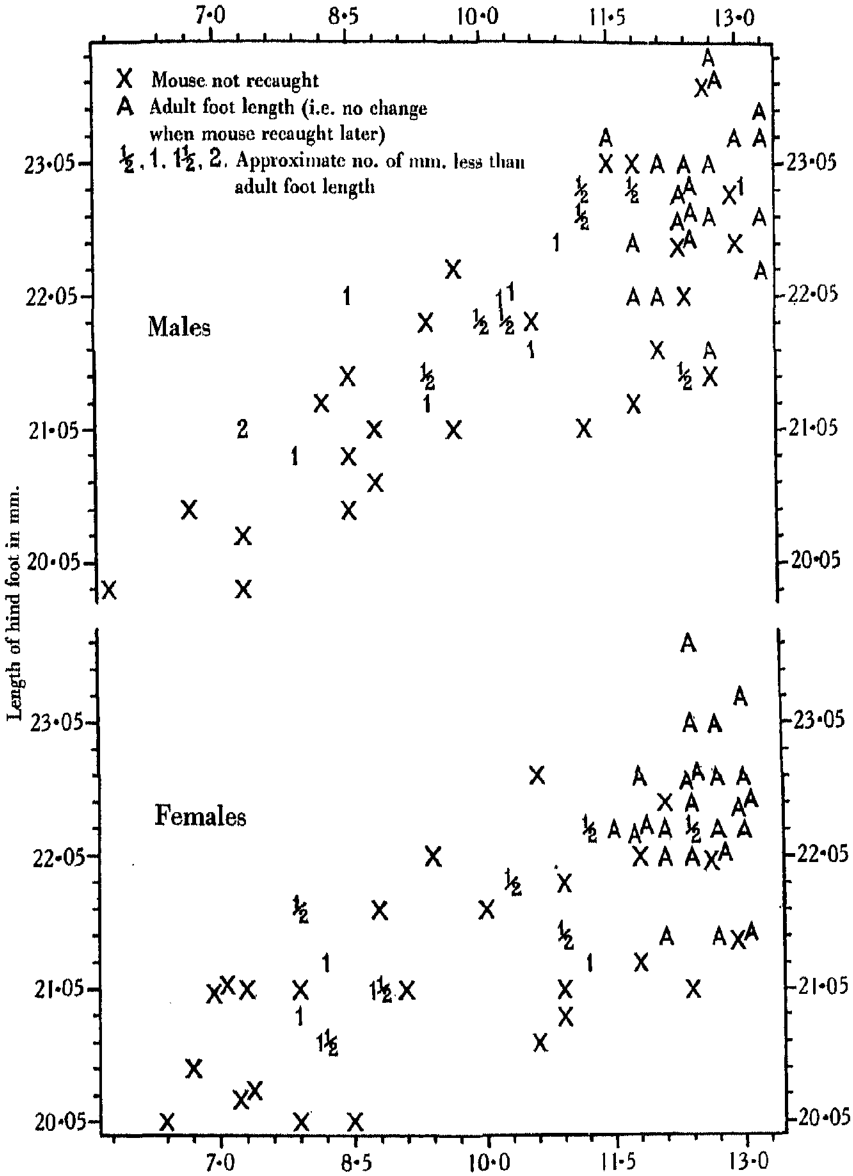


Fig. 3. Young live *Apodemus*, Holwood Park, 1937–8 and 1938–9. To show that length of hind foot rarely increases after a weight of 12·0 g. is reached in males, or 11·5 g. in females.

characteristic of the mouse, and is especially useful in establishing local differences. Fig. 3 shows that the hind foot of the Holwood Park mice ceased to grow after the mice had reached a weight of 12·0–13·0 g., about half the maximum weight of the males. Where

the mouse was caught only once the measurement (marked by a cross) throws no direct light on the final value of the foot length, but it combines with the rest of the data to give a scatter diagram indicating that foot length was still increasing in growing mice at the lower body weights. Where the mouse was caught more than once the measurement is either marked by a letter *A*, indicating adult value, when no subsequent increase was found, or by a number, representing the number of millimetres short of adult value. It is clear that in almost every case where repeated measurements were obtained four or more weeks later (see p. 149), the hind foot had ceased to grow in length by the time the mouse weighed 13·0 g., and very often before that. Thus this measurement gives no indication of the age of any mouse weighing more than 13·0 g.

In an April population (Table 1) the correlation between length of hind foot and length of head and trunk ($r = 0\cdot56$ for males and $r = 0\cdot47$ for females) is very similar to that between length of tail and length of head and trunk; that is to say, it is not sufficiently great to give a clear indication whether an exceptionally small mouse in such a supposedly adult population is likely to be inherently small or merely delayed in growth and therefore to be rejected for statistical purposes.

Table 1 *compared with human data.* The coefficients of variation obtained by Ruger (1933) have already been referred to (p. 141). He gives correlation coefficients for span and weight, stature and weight, sitting height and weight, sitting height and span, ranging from $r = 0\cdot554$ to $r = 0\cdot598$ for men, and from $r = 0\cdot434$ to $r = 0\cdot595$ for women, those for women being quoted from Elderton & Moul (1928); these coefficients are very similar to those for the *Apodemus* material of Table 1 and, as there, are less in every case for females than for males. Much greater correlation is found in man between stature and span ($r = 0\cdot818$ for men, $r = 0\cdot824$ for women) and between many skeletal measurements. The material from which these coefficients were derived, however, was of too different a character from ours to warrant detailed comparison.

A reference to Sumner's (1926) data on American deer-mice is also called for here. Besides other measurements and colorimetric estimates, he gave the means and standard deviations for weight, body length (= head and trunk length), and foot length, in three varieties of the *Peromyscus polionotus* group from Florida and Alabama. These are evidently much smaller mice than our *Apodemus sylvaticus*, and quite differently proportioned. For geographical reasons his data are again not directly comparable with those in Table 1, and comparison is better postponed until we deal with local variation.

Stomach weight. The part which stomach content plays in the total weight of a mouse can be judged from Fig. 4, the data for which came from the series of dead mice trapped with break-back traps for their skeletons (p. 138). The stomachs were cut at the pylorus and at the lower end of the oesophagus, and the spleen removed. As no material difference could be detected between the stomach weights of the two sexes these have been bulked together in the histogram.

It will be seen that the range of weight in 268 mice was from 0·3 to 4·0 g., with a mean of 1·2 g. and a standard deviation of 0·79 g., but that stomachs of 0·3–0·8 g. were the commonest. Weights above 3·0 g. were rare, and the one 4·0 g. stomach, that of a female, was so enormously distended that it appeared to fill the whole abdominal cavity. When the stomachs of two male and two female mice were scraped clean of contents, one of them weighed 0·3 g. and the others approximately 0·25 g., so that the ten stomachs in the 0·3 column in Fig. 4 may be regarded as empty. For younger mice the upper limit of stomach

weight will of course be less, varying with the size of the mouse, while the lower limit may be as little as 0·1 g., as we have found in the unscraped stomach of a 7·0 g. youngster.

The prevalence in the snap traps of mice with empty or nearly empty stomachs indicates that usually only hungry mice seized the bait. Mice living near the site probably found the traps early in their nightly wandering in search of food; they rarely had time to swallow the bait, though they may have eaten some of the oat grains scattered in front of it as an extra attraction.

But the mice with whose growth this paper is mainly concerned were caught alive in large box traps,* to the end of which was fixed a ½ lb. cocoa tin with bedding, two peanuts, and a large pinch of oats. Together with the bait (another peanut), this was more food than they usually ate, so that when weighed these live mice are less likely to have had an empty stomach than those caught in snap traps. On the other hand, with excess of food available, the stomach content may have depended on the time of day at which they were taken from the traps to be weighed, and this varied from 10 a.m. to 7 p.m. according to

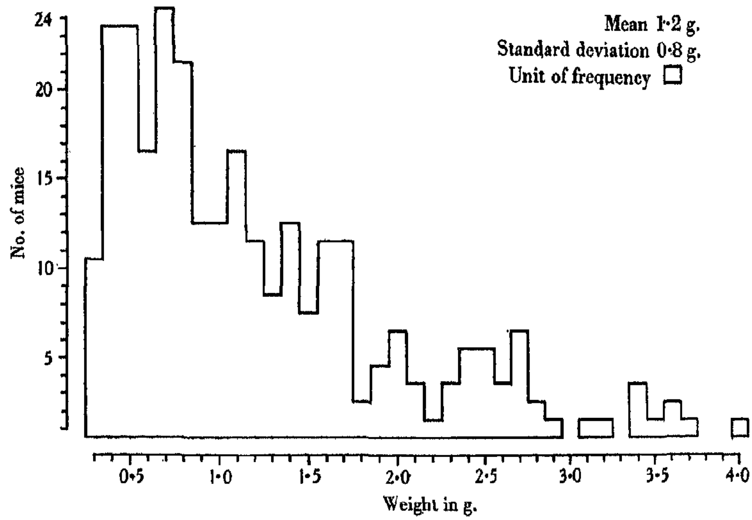


Fig. 4. Frequency distribution of stomach weight in 268 male and female *Apodemus* caught in four localities between 28 March and 13 April 1937, 1938 and 1939. 0·3 g. is the weight of an empty stomach.

the size of the previous night's catch. Elton *et al.* (1931, p. 716) recorded some observations on the times of feeding activity in *Apodemus*, and on the quantities eaten, but more data about this are needed. All we can say here is that a variation of as much as 4·0 g. in the weight of a full-grown mouse *may* be due to stomach content, while any variation of up to 1·5 g. must so commonly be due to this factor that it should be discounted as an indication of change in body weight *where only a single mouse is concerned.*

Reproductive organs. The growth of the reproductive organs in the male, and the presence of embryos in the uterus of the female, are among the factors influencing body weight. We have dissected out and weighed together the testes, vesiculae seminales, Cowper's glands and penis of all dead male *Apodemus*. We have also weighed together in their uterine casing any embryos large enough to make visible swellings in a dead female. Some of these mice

* Selfridge traps (Elton *et al.*, 1931, p. 14). We are indebted to Messrs Selfridge for supplying us with these traps at a reduced rate.

were found dead in the box traps set in Holwood Park, and some were caught in snap traps elsewhere.

(*a*) *Male.* The weight changes of the reproductive organs of a male mouse appear to be correlated with the changes in its body weight. We shall presently show that in midwinter only very small males gain appreciably in body weight, while larger mice are stationary, and autumn adults lose weight (pp. 158–161 and Chart 1, C 1, D 1, D 2). Fig. 5 shows the weight of the reproductive organs of over 200 male mice of varying weights in different

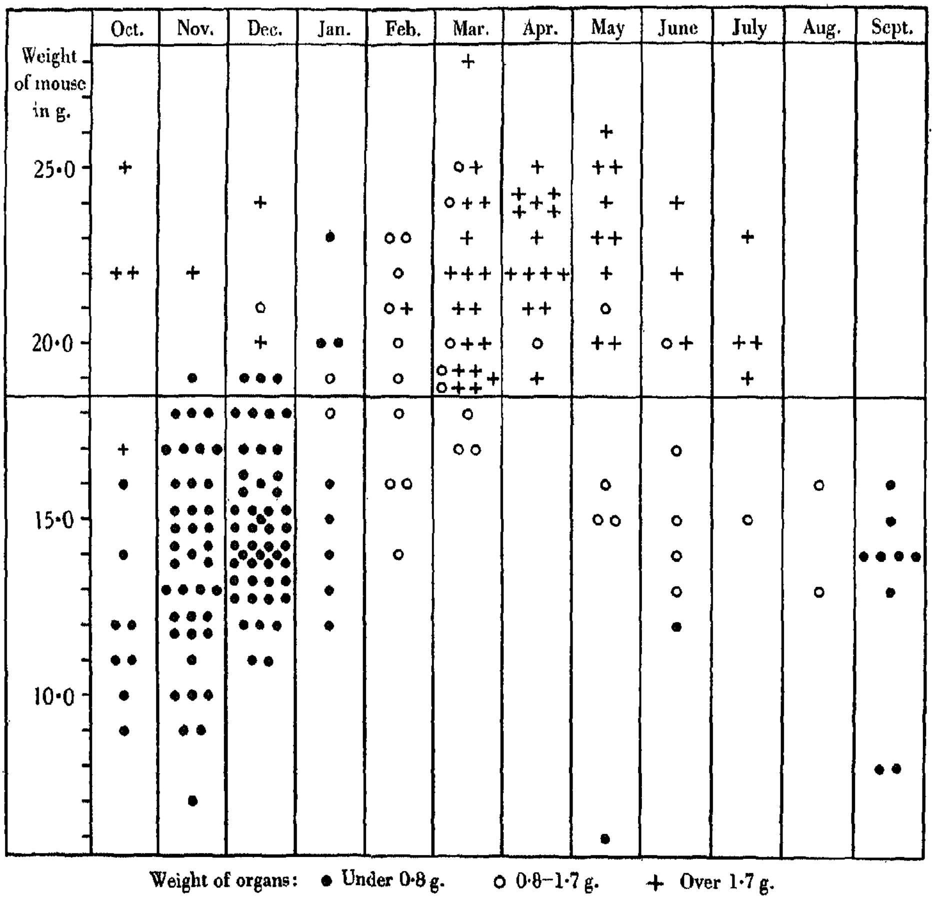


Fig. 5. Weight of reproductive organs in male *Apodemus* of different weights at different times of year. The mice were caught in several years and several localities.

months of the year. Out of 112 of these mice caught in the months September to January inclusive (the majority in November and December), only fifteen weighed more than 18·5 g., the mean weight being 14·5 g. as in the November and December histograms of live mice on p. 159. Only six of these fifteen mice of adult or nearly adult size had well-developed reproductive organs weighing more than 1·7 g., and only two had organs of intermediate weight; while the remaining seven, and all but two of the mice of under 18·5 g., had organs weighing less than 0·8 g., by far the most frequent weights being from 0·1 to 0·3 g. In

several mice weighing from 17·0 to 19·0 g. the testes had evidently been larger, but were withdrawn into the abdomen and appeared like empty, shrunken sacs; the vesiculae seminales and Cowper's glands were also shrunken. These mice may have already reached maturity, but had undergone loss of weight and of reproductive activity at the onset of winter, and were in a condition similar to that described by Brambell & Rowlands (1936) in bank voles.*

After January, when our live records show that the majority of male mice gain in body weight (pp. 154, 155 and Chart 1, C3), it appears that their reproductive organs also grow rapidly in most years, coming into the intermediate group of 0·8–1·7 g. The eleven February mice shown in Fig. 5 were all in this stage except one, in which the organs already weighed 2·0 g. Among other mice in this stage were:

(i) A single December mouse of 20·5 g., whose testes were still unshrunken though small.

(ii) Two advanced January mice.

(iii) Several March mice, presumably slow in their development.

(iv) A group of young summer mice, all under 17·0 g.

(v) Three large mice caught in April, May, and June, of whose significance we are in doubt.

All but one of the 55 mice with reproductive organs weighing over 1·7 g. were themselves over 18·5 g., and therefore of adult or nearly adult size; the majority of the March mice had reached this stage, the weight of their reproductive organs ranging as high as that of adult mice in later months.† Since the range of body weight of adult spring and summer mice is from just over 18·0 to just over 28·0 g., and the range of their reproductive organs is from 1·8 to 2·8 g., if there is any correlation between the two weights it appears that in spring and summer about 10 % of the body weight may be attributed to the reproductive organs. This is very different from what may be expected in midwinter, when these organs contribute only 0·1–0·3 %. We have very little data for any one month in the summer, however, and cannot trace the month to month connexion between the two weights. There is some slight evidence from our records of live mice that a temporary loss in body weight occurs after breeding (p. 155), and it would be of interest to know if this is reflected in the reproductive organs.

For April, a better picture of the range in body weight and in weight of reproductive organs and of the correlation between them ($r=0{\cdot}66$), is given by the scatter diagram in Fig. 6. This represents 149 male mice of which the majority were probably adult; they are among those described on p. 138 as caught in snap traps between 28 March and 12 April in four different localities in 1937, 1938 and 1939.‡

(*b*) *Female*. Of 121 females caught at the same time as the April males just referred to, thirteen had embryos visible in the uteri, indicating that the breeding season had started in all three years, though there were actually some local differences in this respect. These

* We understand that these authors have material for a similar paper on *Apodemus*; much of our own work must be complementary to theirs, and we very much regret that they have not yet been able to publish it.

† Nineteen males were caught in February and March 1942, too late to be included in Fig. 5. In this year both January and February were exceptionally cold months; the mice caught in February corresponded in body weight, and in the development of their reproductive organs, with those caught in January in previous years, while those caught in March corresponded with February mice of previous years.

‡ Since the four groups of mice from these four localities differ among themselves and from the Holwood mice in size (including that early maturing element hind foot length), and therefore in body weight, the rough estimate of 20 g., given on p. 141 as the weight below which most Holwood males may be expected to be still growing, cannot be applied to these data.

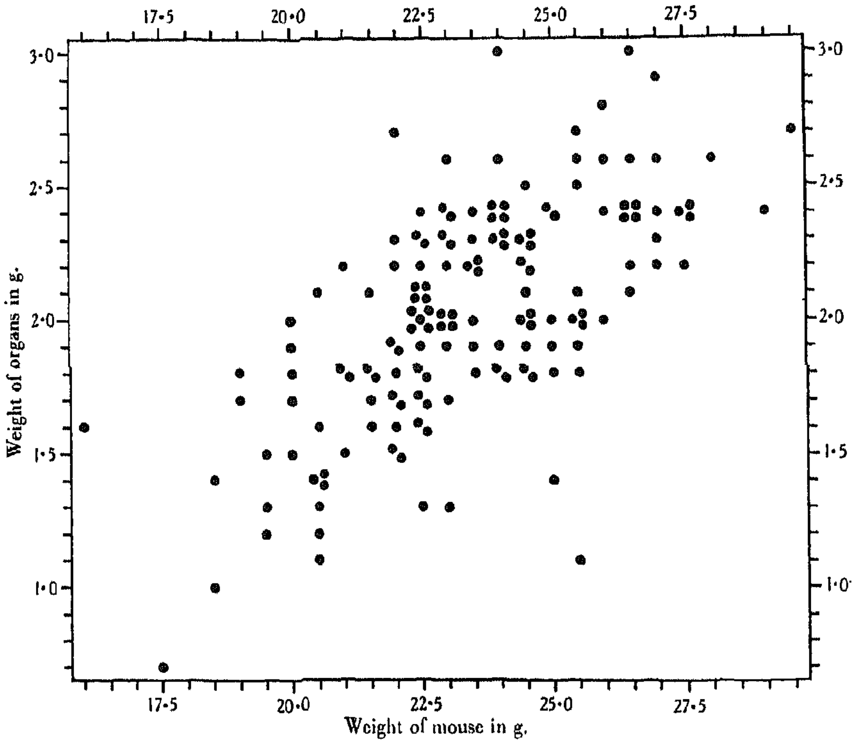


Fig. 6. Correlation between weight of mouse and weight of its reproductive organs in 149 male *Apodemus* caught in four localities between 28 March and 13 April 1937, 1938 and 1939.

Table 2. *Female* Apodemus *caught between* 28 *March and* 13 *April in* 1937, 1938 *and* 1939, *in four different localities. Weight distribution of those with and without embryos.*

Weight of mouse to nearest g.	No. without visible embryos	With visible embryos			Total no. of cases
		No. of cases	No. of embryos	Weight in g. of embryos and uterus	
15·0	5	—	—	—	5
16·0	8	—	—	—	8
17·0	27	—	—	—	27
18·0	16	1	5	0·6	17
19·0	16	1	4	1·1	17
20·0	12	1	4	*S*	13
21·0	14	4	?, 4, ?, 4	0·3, 0·3, *S*, *S*	18
22·0	6	—	—	—	6
23·0	2	4	3, 4, ?, 6	4·5, 1·8, *S*, *S*	6
24·0	2	1	4	1·5	3
⋮					
28·0	—	1	4	6·6	1
	108	13			121

females ranged in weight from 14·5 to 24·5 g., with an outlier at 28·3 g. They are shown in Table 2, which gives particulars of those with visible embryos; where the letter *S* is recorded in the last column but one the weight for embryos and uterus was negligibly small compared with the body weight.

It is clear from these data that even at the beginning of the breeding season, heaviness in a female is not necessarily due to the weight of embryos. In our live trapping in 1939 we found that there was a period of universal weight increase for females between March and April, when mice of all weights showed a wide range of increase (p. 158, and Chart 2, C5, C5X, D5, D5X), although the weather was very cold and no young appeared in the traps until May: there seemed no connexion between pregnancy and previous weight. Prof. Brambell's material may throw more light on this point. In non-pregnant females the reproductive system contributes but an insignificant amount to the weight of the mouse.

4. Monthly weight records

(*a*) *General discussion*

Our best records of weight changes in live mice are from the season 1938–9. We trapped each month in two neighbouring parts of the Holwood Park woods,* a week in one part followed by a week in the other, from the beginning of the third week in November until the summer breeding season.

We have given the weeks of the year arbitrary numbers, so that one year may be the more easily compared with another. Week 1 always starts on the first Monday in October. This has the disadvantage that there may be as much as six days' difference between a week in one year and the week with corresponding number in another year.

Month	West of footpath		East of footpath	
	No. of week	Dates	No. of week	Dates
November	7	14th to 17th	8	21st to 25th
December	11	12th to 15th	12	19th to 23rd†
January	15	9th to 12th	16	16th to 19th
February	19	6th to 9th	20	13th to 16th
March	23	6th to 11th	24	13th to 18th
April	29	17th to 22nd	30	24th to 29th
May	33	15th to 18th	—	—
June	—	—	38	19th to 22nd
July	43	26th to 27th	—	—

It will be seen from this list that trapping was repeated every four weeks until after the March catch, but that six weeks elapsed between the beginning of the March catch and the April catch, five weeks from May to June, and five weeks from June to July. In the first four months we usually trapped for four nights only in each area, the first four nights of each trapping week; in March and April for six nights in each trapping week, in May and June for four nights, in July for only two nights. In each of the last three months we trapped in only one part of the woods in order to interfere less with breeding.

Chitty (1937, p. 43) has shown that mice caught one night and set free next day are

* We are greatly indebted to the late Lord Stanley and to Lady Stanley for permission to trap on the Holwood estate at Keston in Kent, the site of a great Iron Age camp and once the home of William Pitt, who here enclosed many acres of common land, tended the woodlands, and planted trees on the bulwarks.

† All traps were taken in on the night of 21 December owing to a fall of snow.

likely to reappear in the same or a neighbouring trap night after night. This keeps other mice out of the traps, or, if surplus traps are set, greatly adds to each day's labour. We therefore placed all the mice caught in cages, a separate cage for each trapping site, and released them in their home area only after the traps had been taken up at the end of the week's trapping. An exception was made in the case of pregnant or lactating females, which were set free each day; even these generally came back to the trap next night, and we are afraid that some interference with breeding must have taken place. At times young were born in the traps, or before the mother could be set free, and she then generally killed and ate them, or else left them to die. Once when we set a mother free with her young where caught, she did not take them back to the nest, but they were found abandoned and dead on the same spot next morning. At other times a released female would show a great drop in weight when recaught next day, suggesting that her family had been born in the interval (p. 153 and Fig. 8, group VI, no. 600); as she may have been away from the nest for more than 12 hrs. it is all too likely that the nest young died meanwhile.*

Even out of the breeding season, during a period of active growth, there seems some evidence that frequent catching interferes with weight increase. Mice caught in two consecutive weeks, as happened frequently to those living on the adjoining borders of our two areas, often weighed less at the second catching. We have further evidence of this from other years when one area was trapped again after a short interval. It is conceivable that the commencement of breeding in a whole population might be delayed by too frequent trapping and too long periods of captivity.

Chitty (1937, p. 46) has given a list of eight disturbing effects of trapping, and our method of keeping the mice together in cages for several days is likely to enhance and add to these. On the other hand, it probably enabled us to catch a greater number of mice and thus to have better data for statistical treatment; for we have evidence that the local inhabitants can be removed from an area in a night or two, and that outsiders then wander into it, if suitable weather conditions prevail. This evidence we hope to give in a future paper on wandering.

We caught 343 different mice in the season 1938–9, of which 134 were never recaught, including nine which died in the traps or cages. Twenty-three mice were under 10·0 g. in November and December, but only nine of these were ever caught again, and only four reappeared each month throughout the season. These four mice gave us our most complete individual weight records, which are graphically represented in Figs. 7 and 8, groups I–VI, together with those of a number of other mice of greater weight when first caught. In these graphs as many mice have been selected as could be represented clearly in a few text-figures. Mice of less than 7·0 g. rarely appear in traps, and it is probable that they do not wander out from the nest until about this size unless the nest is destroyed or their mother fails to return to it; the smallest mice we have ever caught in traps were two males of 5·5 and 5·8 g.

Besides these fairly complete individual records we had many lasting over shorter periods. The spot diagrams of month to month weight changes (Charts 1–3) and the monthly weight frequency histograms of Figs. 10 and 11, are an attempt to utilize these data by massing them, which should to some extent smooth out hour to hour variability due to such causes as change in stomach content.

* Burt (1940, p. 15) has, however, 'one record of newly born *Peromyscus* which lived sixty hours without parental care of any kind'.

(*b*) *Individual weight records. Growth graphs*

Males. Fig. 7, group I, shows the rapid rise in weight of four young males between the November and December catches. The rise was greatest in the smallest mouse, but not sufficient for it to overtake the others. Between December and January the rise was very slight; in the smallest mouse, no. 468, it was still slightly greater than in nos. 423 and 528,

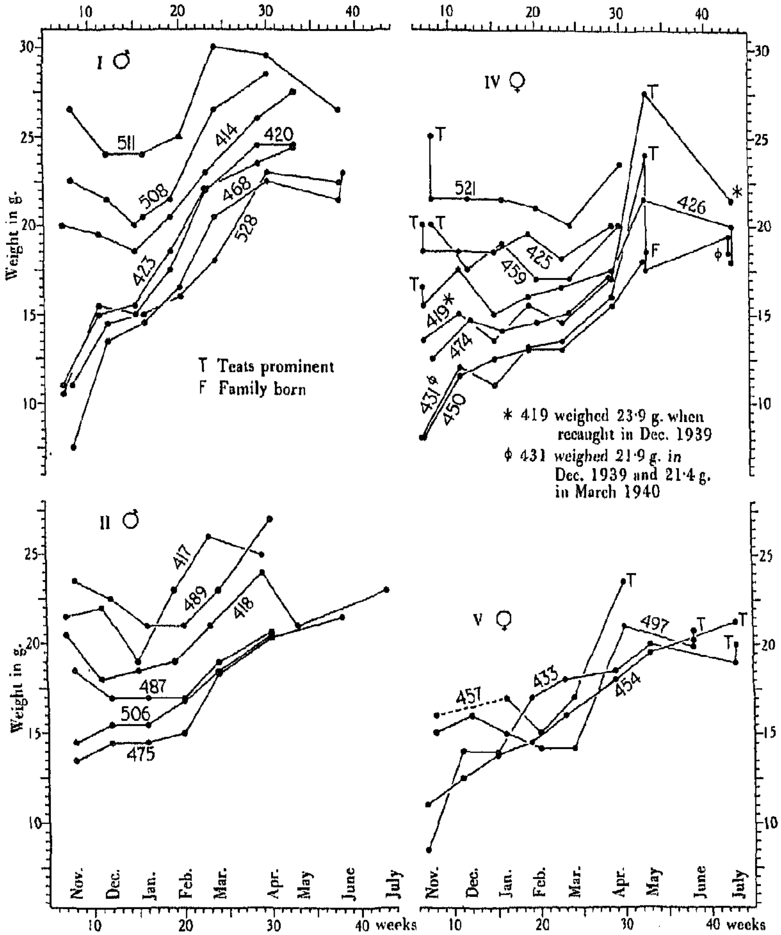


Fig. 7. Individual weight records of live *Apodemus* caught at monthly intervals in Holwood Park, November 1938 to July 1939.

whereas no. 420 actually lost a little. At the top of group I, nos. 508 and 511, and possibly also no. 414, were already of adult size in November; they all showed loss of weight between November and December, and the first two continued to drop until January, while no. 511 remained stationary. All the mice began to put on weight after the January catch, the older mice still remaining ahead of the younger mice when the breeding season was reached. Taking them all together the period of most rapid growth was between February and March; the increase tended to slacken after March or April, and was followed in three of them by actual loss in weight.

Fig. 7, group II, shows three more mice of adult size losing weight between November and January: nos. 489, 417 and 418. No. 487, of only 18·5 g. in November, also lost weight, and may also have reached maturity in the autumn breeding season, but if so it was an exceptionally small mouse. Nos. 506 and 475, of 14·5 and 13·5 g. in November, were evidently older than the young mice in group I and showed a much smaller increase than these by December; they did not change between December and January, and again showed smaller increases in the spring, both reaching only 20·5 g. by April, whereas the younger mice in group I reached from 22·5 to 24·5 g. The older mice in group II, like those in group I, kept far ahead of the younger ones, with the exception of no. 487 which was overtaken by them. This may be another indication that no. 487 was an exceptionally small mouse; it had a very short hind foot, only 22·0 mm., but, as has already been shown,

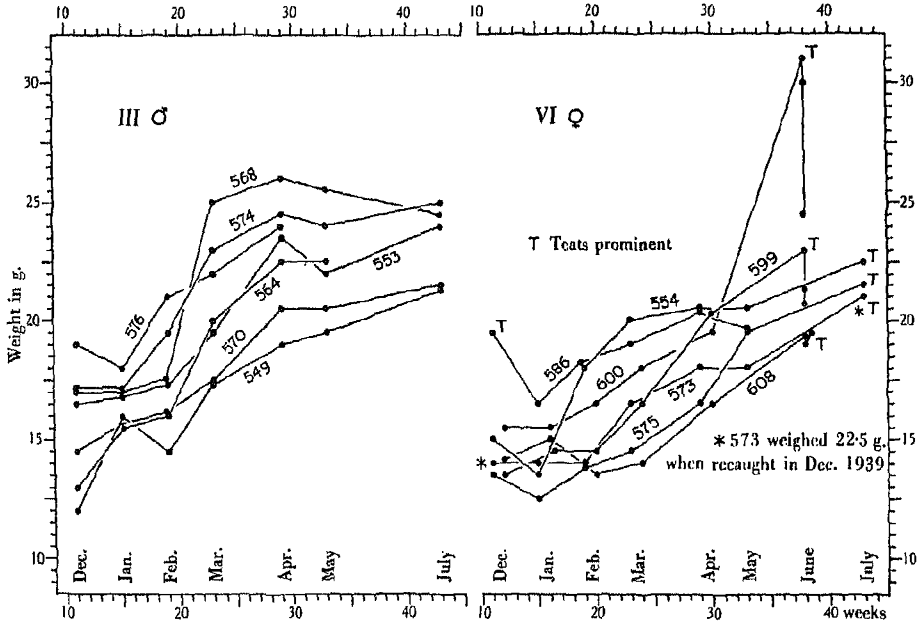


Fig. 8. Individual weight records of live *Apodemus* caught at monthly intervals in Holwood Park, December 1938 to July 1939.

there is not a very high correlation in adults between length of hind foot and size of mouse. That there is also no close association between foot length and rate of spring increase in weight can be seen in Fig. 9, where, however, one of the two dots in the bottom left-hand corner represents no. 487.

Fig. 8, group III, contains the weight graphs of seven male mice which were not caught until December. A comparison with Fig. 7, group I, suggests that nos. 564 and 549 were very young mice in November, as they showed a sharp rise in weight between December and January. The larger mice, nos. 553, 568 and 574, remained nearly stationary, and judging from other mice not included in the diagrams were likely to have had about the same November weight of 16·5–17·0 g., though there were also a few smaller mice in November which increased to that weight by December and January (Chart 1, D1, D2). The largest

mouse of all, no. 576, lost weight between December and January. They all showed the rapid increase which was common to male mice in the spring, and which slackened off sometimes at the end of March, sometimes in April.

Females. Fig. 7, groups IV and V and Fig. 8, group VI, give the weight graphs of a number of female mice. They show that young females soon fell behind young males of comparable November weight; no. 433 was an exception, keeping up with the males until after the February catch. There is much irregularity and overlapping among the females, and no general rule is apparent. Losses between December and January were frequent among both smaller and larger mice, while the three November adults, nos. 521, 425 and 459, still showed losses between January and March. Many even of the younger females showed no marked spring increase until after the March catch, though others started increasing between February and March, and a few between January and February. Between March and April all showed gains, some of which may have been due to early pregnancy; they were comparable in amount with those of the males between these months, and between January and February, but were not as great as those of the males between February and March.

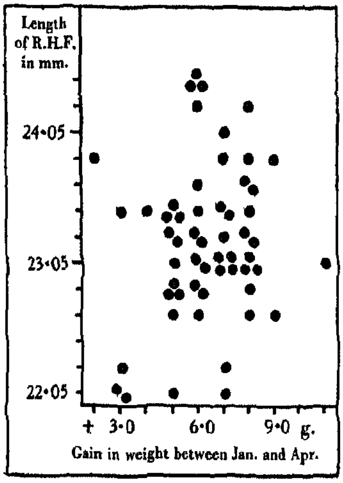


Fig. 9. Live male *Apodemus*, Holwood Park, 1939. To show that there is almost no correlation between length of hind foot and rate of spring increase in weight.

Advanced stages of pregnancy may be picked out by a sudden sharp rise in the graph, while a long vertical fall following this in the same week may be presumed to indicate the birth of young, though we had no other evidence of this when it occurred during release from captivity (see p. 150). Two examples may be taken from the graphs: no. 431, group IV and no. 600, group VI. No. 431 weighed 24·2 g. on 16 May. She was released near the same trap that afternoon, and found again next morning in another trap 10 yards off, together with three young which she had killed; her weight on this second day was only 18·6 g. She may have given birth to more young and eaten them, but in case there were still young unborn she was released again in the afternoon. Next day, 18 May, she was recaught in the original trap and then weighed only 17·4 g. No. 600 was caught on four consecutive days in June; she had the unusually heavy weight of 31·0 g. on the first day, had lost only a gram on the second day, but on the third day had dropped to 24·7 g., a weight which she retained on the fourth day. Apart from this drop in weight there was in this case nothing to indicate the birth of young.

Summary of information from individual weight graphs

Males. (i) Between November and December the smallest males gained most weight, larger mice remained stationary, while the largest lost weight.

(ii) Between December and January some of the smallest mice still gained a little weight, larger mice remained stationary or continued to lose weight.

(iii) Nearly all males put on weight after the January catch.

(iv) The period of most rapid male increase was between February and March.

(v) After March or April gains were smaller and some losses occurred.

(vi) Most of the older mice were still ahead of most of the younger mice at the beginning of the breeding season.

Females. (vii) Between November and December the weight changes of the females were similar to those of the males, except that the smallest females did not gain so much weight as the smallest males.

(viii) From December to March no general rule was apparent.

(ix) Between March and April increases were universal.

(x) In this year the breeding season started at the end of April.

(c) *Combined weight records. Spot diagrams of month to month weight changes*

The information just summarized comes from the very small sample of the population represented in the individual weight graphs. It is confirmed and amplified by the spot diagrams of month to month weight changes, which represent rather larger samples of the 1938–9 population. There are three series of these diagrams shown on Charts 1–3, and it is to these that the letters and numbers in the present section refer; each puts on record information not given by the others, and between them they enable most of the data from all the mice caught more than once to be utilized.

A series. Weight in each month plotted against loss or gain by the following month.*

C series. Weight in November plotted against loss or gain in each succeeding monthly interval.

D series. Weight in November plotted against weight in each succeeding month.

The CX and DX diagrams are given because we had a very large catch in January and the records of a number of new mice start then. The records of those males caught in November seem to show that autumn adults could still in January be distinguished by weight from autumn young. Judging from the data given in Chart 1, D 2, the former might be expected not to have fallen to a weight of less than about 18·5 g. and the latter not to have increased to this weight. For adult females it is convenient to take a lower January limit of 17·5 g., for though some had undoubtedly dropped below this, it was then impossible to distinguish them by weight from growing young.

Adult males' winter loss of weight and early spring increase. The bulk of the November males of 18·5 g. and over lost weight by December, some of them as much as 2·5 g. (Chart 1, D 1, C 1).† Between December and January some again lost weight and some remained stationary (Chart 1, C 2), the bulk of them in January being from 1·5 to 2·5 g. behind their November weight (Chart 1, D 2). Between January and February some began to put on weight again, but some were stationary (Chart 1, C 3, C 3X) so that while a few were now ahead of their November weight, others were still from 1·5 to 2·5 g. behind it (Chart 1, D 3). Between February and March nearly all showed big gains, most commonly of about 2·0 to 3·0 g. (Chart 2, C 4, C 4X); all now reached their November weight again, and some were as much as 3·5 g. ahead of it (Chart 2, D 4). Between March and April the gains were very irregular and there were two losses, the range of increase being from − 0·5 to + 4·5 g. (Chart 2, C 5). The bulk of the adult mice were now from 2·5 to 3·5 g. ahead of their November weight; the heavier the mouse in November the heavier it usually was in April, the range of weight being then from 23·5 to 30·0 g. (Chart 2, D 5).

* A1 = C1; diagrams A2, 3 and 4 have been omitted for economy in space.

† The weights given in the text are the actual weights of the mice correct to 0·1 g. In the spot diagrams the data are grouped into 1·0 or 0·5 g. divisions, and so may sometimes appear not to correspond with the text.

Young males' winter and early spring changes. Of the November male young (under 17·5 g.), those under 15·5 g. had all gained weight by December, and, on the whole, the smaller they were the more they had gained; those under 11·5 g. had advanced very fast indeed, some gaining 5·0 to 6·0 g. and overtaking many of those above this weight which showed only small increases (Chart 1, D 1). Between December and January the two mice which were smallest in November had both gained another gram, whereas the slightly larger mice had gained less than this or lost a little (Chart 1, C 2); for the whole period, November to January, it remained true that the smaller the mouse in November the more it was likely to have gained (Chart 1, D 2). Between January and February the bulk of the November young put on weight, some as much as 4·0 g. (Chart 1, C 3), but the most rapid period of spring growth in this year, for the young as for the adults, was between February and March, when there were no losses in weight among the younger males but a wide range of increase, gains of 4·0 to 4·5 g. now being very common (Chart C 4, C 4X). Six weeks later, towards the end of April and just before the breeding season, some further big increases were shown, but on the whole the rate seems to have slowed down, the commonest gain being about 2·0 g. (Chart 2, C 5, C 5X); it still held good that the smaller a young mouse was in November the greater was its expected gain, the smaller young having by now overtaken the larger young within the range of 20·5–24·5 g. (Chart 2, D 5).

An overlap in weight by February. Between the November young and the November adults an overlap in weight had already started in February, just before the period of maximum spring growth; this overlap was then between 18·5 and 20·5 g. (Chart 1, D 3, D 3X), while in March it was between 20·5 and 24·5 g. (Chart 2, D 4, D 4X), and in April between 21·5 and 26·5 g. (Chart 2, D 5, D 5X). From February on, therefore, there was no clear division by weight of the November male adults from the November male young, though the bulk of the former were still heavier than the bulk of the latter.

Male weight losses in the breeding season. After the start of the breeding season our trapping was confined to one only of our areas in May, to the other in June, and to the first again, but for two nights only, in July. This meant that we caught many less mice in each month, and among these there were very few of our November catch: in May only three of the November male adults and four of the young, in June two of the adults and nine of the young, in July one of the adults and one of the young.

In May and June the bulk of the November young weighed from 20·5 to 24·5 g., their average weight in May being 22·7 g., a gram higher than in June. The few remaining autumn adults were very scattered in weight, overlapping the young in range but with an average still a little ahead of them (Chart 3, D 6X, D 7X). Between April and May there were nearly half as many losses as gains, and although these had apparently no relation to the weight of the mouse in November (Chart 3, C 6X), the losses were confined to mice which in April were 21·5 g. and over, while the laggards then under 21·0 g. all showed gains (Chart 3, A 6); in all, nine mice lost weight, nineteen gained weight, and four remained the same. Between April and June losses outnumbered gains by fifteen to twelve (Chart 3, A 7). It is perhaps significant that all these losses occurred in the early part of the breeding season, during the two months in which pregnant females first appeared in the traps. The net result, a male population almost stationary in weight, is best seen in the weight histograms (Fig. 10).

A possible summer period of further weight increase. Of the thirty-four males caught in May only thirteen were recaught during the two nights when we trapped over the same

area in July. This is a very small number from which to draw conclusions, but since eleven of the thirteen had gained from 1·0 to 2·5 g., and only two had lost weight, it is possible that a new period of weight increase had followed the period of losses (Chart 3, A 8). The increase in weight of this very small sample of the July population is again best seen in Fig. 10; leaving out the new season's young, the average weight was 24·3 g., whereas the May and June averages were only 22·9 and 22·8 g.* It may be noted here that no young were born in the traps or cages in July, and that none of the females caught appeared to be in an advanced state of pregnancy, the heaviest weighing only 22·3 g.

Table 3. *Adult females' loss of weight after capture*, 1938–9

	424	425	422	415	416		488	495	476	465	516	466	521
14 Nov.	—	—	23·7	25·7[2]	30·6[2]	21 Nov.	—	—	21·5	22·9	—	24·8	—
15	16·6	20·0	—	—	—	22	17·6	20·6	21·7	21·8	—	—[3]	—
16	14·2	18·5	22·1	—	—	23	18·8	21·5	—	21·9	23·0	—	—
17	16·0	—	18·3[1]	—	—	24	15·9	19·2	18·8	19·9	22·1	—	24·8
	—	—	—	—	—	25	16·8	19·4	20·6	21·3	19·9	—	21·6
	—	—	23·0	18·3.	26·5		—	—	—	—	—	18·6	—

	715	743	452	620	431	717		768	643	599	637	660	600
15 May	—	—	—	23·2[4]	—	—	19 June	22·3	23·0	23·1	24·7	26·5	31·1[7]
16	20·0	20·9	22·2	—	24·2	26·3	20	19·7	21·3	21·6	22·4	20·0	30·0
17	19·8	20·5	20·5	17·6	18·6[5]	24·5	21	—	19·9	21·1	19·4	—	24·7
18	19·2	18·3	18·6	—	17·4	19·5[6]	22	—	20·2	20·3	—	—	24·7

[1] 422. Ill, kept in for two days; weight increased again to 23·0 g.

[2] 415 and 416. Both appeared nearly parturient and were not released. Young subsequently heard in cage, but died and were eaten. Mothers weighed again on 24 November and released.

[3] 466. Vaginal discharge on 22 November; not weighed, kept in cage. 24 and 26 November, young heard. 1 December, six dead young seen, mother weighed 18·6 g.

[4] 620. Appeared nearly parturient, kept in. Young seen and heard next day; some killed and partly eaten by mother, the rest placed by us in some hay with the mother in the woods where she had been caught. Young, all dead, found still in hay on third day; mother recaught weighing 17·6 g.

[5] 431. Three dead young in trap.

[6] 717. Bleeding at vulva.

[7] 600. This very heavy mouse had a very long right hind foot.

The remaining weight losses all occurred while the mice were set free overnight.

Adult females in winter. Among the females the November adults were not so clearly marked off by weight from the November young as among the males. Teats were prominent in November in all female mice of 17·5 g. and over, and also in two out of three of about 16·5 g. A heavy female rarely dropped in weight to less than 17·5 g. when known or supposed to have given birth to a family (p. 150, and Table 3); this is therefore a convenient weight above which to regard all females as adult, and 'November adults' of 1938–9 will here mean those shown on the graphs as 17·5 g. and over in November, though it is probable that females of a lower weight than this do become pregnant.

* The averages given on p. 155 were for the November young only.

Late autumn pregnancies. Among these adults the biggest losses between November and December were in two mice which fell from 24·8 to 21·5 and to 17·5 g. respectively (Chart 1, D 1). The weight graph of one of these has been seen in Fig. 7, group IV (no. 521). Most of the loss had already occurred when the mice were weighed for a second time in November (see Table 3)—the first mouse (no. 521) when recaught after a night's release from captivity, the second (no. 466) at the end of a week's captivity after giving birth to a family of at least six young. Such losses are comparable with those in May when it was known or suspected that pregnant females had given birth to young (p. 158). Three females of about 20·0 g. in November lost from 1·5 o 2·5 g. by December, while one of 17·5 g. remained at this weight (Chart 1, C 1, D 1); these mice may have been nursing mothers in November (p. 161).

Among females, no conspicuous weight increase until after March. Between December and January one mouse in this group gained weight (but not as much as it had previously lost), three lost weight, and two remained the same (Chart 1, C 2); between January and February one gained a little, two lost weight and two remained the same (Chart 1, C 3, C 3X); between February and March, the most rapid period of male growth, three lost weight and two remained the same (Chart 2, C 4, C 4X). The net result was that by March the five surviving November adults were all from 2·0 to 5·0 g. behind their November weight, in great contrast to the adult males which had all by now overtaken that weight, some showing gains of up to 3·5 g. (Chart 2, D 4). In the six weeks between the March and April trappings a period of growth at last set in for all females, and the five surviving adults showed increases of up to 4·0 g. (Chart 2, C 5, C 5X), all but one of them having now nearly caught up with their November weight (Chart 2, D 5). After April only one of these adults was ever caught again—in June, when its teats were prominent and its weight was 26·2 g.

Young females compared with young males in winter and early spring. Among the November young only three very small females of 8·0, 8·1 and 8·7 g. in November put on much weight between November and December, and of these only the largest put on as much as the males of similar size (Chart 1, C 1, D 1). The individual graphs of these three were given in Fig. 7, groups IV and V (nos. 431, 450 and 433). Young females of 10·5–13·0 g. in November did increase slightly, but not nearly as much as the males of this weight, while from 13·5 to 17·5 g. there were small changes in both directions. Between December and January the very young mice ceased their rapid increase, two of them actually losing weight, while those of intermediate size showed some gains and some losses (Chart 1, C 2); young females behaved very like young males in this month, the juvenile population as a whole remaining stationary (Chart 1, C 2). Between January and February the very youngest females were once more putting on weight, but from 10·5 g. upwards losses and gains about balanced one another; this was in contrast to the males, only two of which showed loss in weight, while there was a high proportion of big increases (Chart 1, C 3, C 3X). Between February and March the contrast was still greater; there were fewer losses among the young females but no very big increases, and four remained the same (Chart 2, C 4, C 4X). The net result was that the mice which were smaller in November had gained from 5·0 to 9·0 g. by March while the rest showed intermediate gains and losses in rough accordance with their November weight (Chart 2, D 4). Already since February some of the November young had overtaken some of the November adults within the range of 16·5–18·5 g. (Chart 1, D 3X and Chart 2, D 4X).

The six weeks between the commencement of the March and the April trappings were, as stated, a period of growth for all females. Young mice as well as old showed gains of from 0·5 to 4·0 g., about 2·0 g. being the commonest (Chart 3, A 5 and Chart 2, C5, C5X). These gains were much the same as those of the males for this period, since the male increase had by now slowed down, but nevertheless the bulk of the females were still far behind males of similar age (Chart 2, D 5, D 5X) and, in spite of further male losses in May and June, never overtook them unless in advanced pregnancy (Chart 3, D 6X, D 7X).

Spring pregnancies. As already seen in November, advanced pregnancy reveals itself in these diagrams by a wide gap between an individual and the bulk of the female population,* just as it is disclosed in the individual weight graph by a sudden sharp rise (p. 153, and see also Table 3, p. 156). There were already two such individuals in April (Chart 2, C5, C5X), mice which had weighed only 15·2 and 16·2 g. in November (Fig. 7, group V, nos. 497 and 457). In May four other females were widely marked off from the rest by exceptionally big increases of from 8·5 to 11·0 g. (Chart 3, A 6); the individual graphs of two of these are seen in Fig. 7, group IV (nos. 431 and 419). All four had prominent teats, and two gave birth to young before it was possible to release them; a third (no. 717), after being released twice, was found to be bleeding at the vulva when caught for a third time, and as she had dropped in weight from 26·3 to 24·5 to 19·5 g., may be presumed to have given birth to a family during her second night of freedom. In June there were two other exceptional increases, of 11·5 and 8·5 g. (Chart 3, A 7); after release and recapture the weight of these mice was found to have dropped from 31·1 to 24·7 g. (no. 600), and from 26·5 to 20·0 g. (no. 660), so that they also may be supposed to have given birth to young in the interval; the individual graph of the first of these two is shown in Fig. 8, group VI (no. 600, see also p. 153).† The next largest increase between April and June was of 5·5 g., but in this mouse the drop in weight on recapture was gradual and not easy to interpret; it was from 23·0 to 21·3 g. after the first night, and from this to about 20·0 g. after the second and third nights (Table 3, p. 156, no. 643).

In June and July prominent teats were noted in all but two of the thirty-one females weighing 17·5 g. and over, whereas in May, out of twenty-five such mice, they were noted in only four besides those four already mentioned as in advanced pregnancy.

Apart from these cases of pregnancy, the bulk of the young females showed advances of from 1·5 to 3·5 g. between both April and May, and April and June; two remained the same in May and two had lost weight (Chart 3, C6X, C7X). The only female which lost weight between the April and June catching was one of the two which had put on so much weight between March and April as to have suggested an early pregnancy (see above, and Fig. 7, group V, no. 497).

None of the females caught as adults in November, or presumed to have been adults

* It must not be forgotten that any *single* exceptional measurement might be due to an exceptionally full stomach (p. 144 and Fig. 4) or to an error in recording. Single errors may also creep in in some such way as follows. In March 1941 a female mouse of 29·2 g. was recorded on capture as appearing to be in an advanced state of pregnancy. She was not released and next day was found dead in the cage, and on dissection proved to have a mass of tapeworm cysts attached to the small remaining fragments of the liver, and occupying most of the greatly distended abdomen. This mass weighed 13·1 g., which, subtracted from the apparent weight of the female, would leave her true weight, without liver, at 16·1 g., that of a not yet fully grown mouse; when first caught in December she had the adult weight of 19·5 g. Had she not died she would have remained in the records as a unique case of pregnancy in the middle of March.

† This mouse, no. 600, was apparently an exceptionally large individual. The length of its right hind foot was bout 24·0 mm., much above that of the average female.

then by their weight in January, were recaught in May or July, and only two* in June (Chart 3, D7X).

(d) Combined weight records. Monthly histograms

In Figs. 10 and 11 are frequency histograms† showing the monthly weight distribution of the Holwood Park 1938–9 male and female populations. Use is here made of all the mice caught, even those caught only once. In Fig. 12 these histograms are used as a background to a similar series for 1937–8. In that year, instead of an intensive fortnight's trapping at

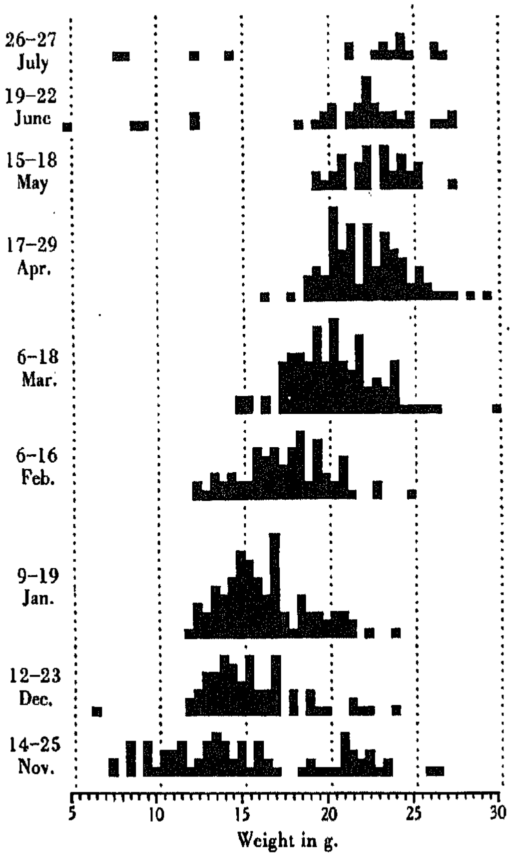


Fig. 10. Monthly weight frequencies of live male *Apodemus*, Holwood Park, 1938–9.

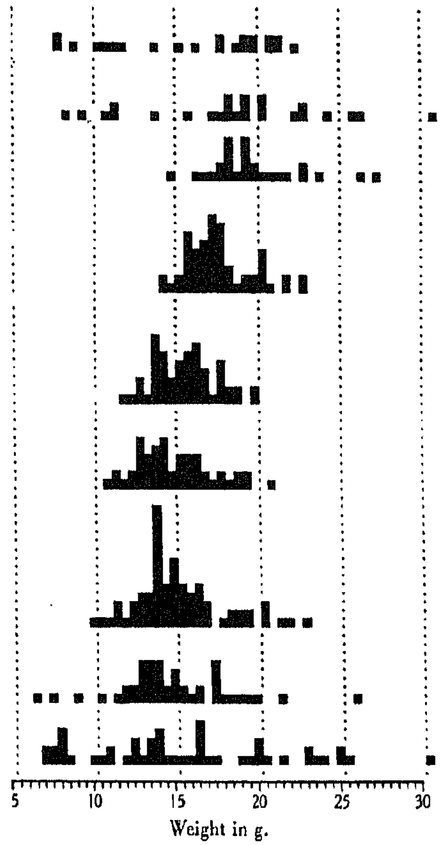


Fig. 11. Monthly weight frequencies of live female *Apodemus*, Holwood Park, 1938–9.

4-weekly intervals, each month's catch was spread over the whole month; we trapped each week, covered a wider area, and shifted our trapping sites according to a different plan. Individual mice were caught at much less regular intervals and so gave less satisfactory records, while the monthly grouping is not strictly comparable with that for 1938–9. The general picture of weight increase in the two years is the same, however, and presents in another way information already given by the individual weight graphs and the month

* One, no. 637, was not caught in April, so is not included in Chart 3, C7X.

† In these diagrams and in Fig. 12 the unit of frequency on the vertical scale equals the distance between the dots on the vertical lines.

to month spot diagrams: a late autumn population reducing its wide weight range by the growth of the small mice and the weight losses of the large mice; a nearly stationary mid-winter population; and a rapid early spring increase among the males followed by a later, lesser increase among the females. Reduced trapping during the breeding season of 1939 extended the picture for that year, and showed a slowing down of weight increase and the reappearance of a wide weight range. Comparison of Figs. 10 and 11 shows once more that in 1938–9 the female population kept up with the male until after January; between February and March was the period of most rapid male increase, whereas the female population did not as a whole gain much weight until after the March catch.

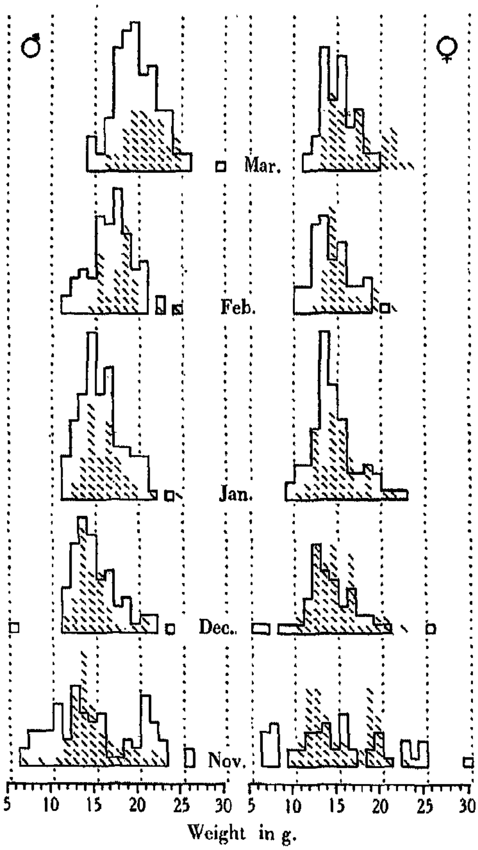


Fig. 12. Monthly weight frequencies of live *Apodemus*, Holwood Park. Plain outline, 1938–9. Hatched, 1937–8.

Where differences between the two years can be detected they can probably be attributed to the marked weather differences. The more obvious differences in the histograms are as follows:

(i) More young of both sexes in November 1938–9.

(ii) More males of over 20·0 g. and females of over 22·0 g. in November 1938–9.

(iii) A group of smaller mice of both sexes in February 1938–9, not present in 1937–8.

(iv) A group of females of over 20·0 g. in March 1937–8, not present in 1938–9.

With regard to these four points, the following comments may be made:

(i) The autumn of 1938 was exceptionally mild and breeding continued late. A number of 7·0–10·0 g. youngsters were caught in November and some even in December (cf. p. 150). All the females of over 17·0 g. and two of 16·5 g. had prominent teats; some were probably nursing young in the nest, while others may have been pregnant (cf. pp. 156–7). All these females were set free each day but they were often caught three nights running and we are afraid that some of the nest young may have died meanwhile (cf. p. 150 and footnote); if so, this would have reduced the number of 7·0 to 10·0 g. youngsters in the December population.

(ii) The population of November 1938 differed from that of November 1937, not only because of the number of very young mice but also because of a greater number of heavy adults. The male group of over 20·0 g. is especially conspicuous in the histogram of Fig. 10, and in weight distribution resembles on a small scale the total male population in April, the beginning of the 1939 breeding season. Most of this group, like the similar group of females, may still have been breeding, but more information is needed about the connexion between weight, season and fecundity in *Apodemus*, which we believe has already been collected by Brambell and Rowlands (pp. 146–7). In the November 1937 population there were far fewer males of over 18·0 g. in proportion to the intermediate group of 13·0–18·0 g. The earlier onset of cold weather is likely to have checked breeding, and the consequent recession of the male reproductive organs to have been accompanied by a loss in weight. The November catch in this year indeed appears comparable with that of December 1938, when adult males had lost from anything up to 2·6 g. of their November weight (p. 154 and Chart 1, C1, D1).

As regards adult females, it is true that in November 1937 we caught at least ten which still had prominent teats, but though eight of these weighed over 18·0 g. and are part of a conspicuous group at the tail of the histogram, there were none comparable with the 1938 group of over 22·0 g. Like the rarity of youngsters of under 10·0 g., and the evidence from the adult males, this again points to the 1937 breeding season having ended before November. One unique little mouse must not be forgotten however: a male of only 7·2 g. caught on 2 December of that year.

(iii) The next conspicuous difference in the histograms is in February. While we can be fairly certain that the differences in November are connected with the long mild autumn of 1938, the reason for this February difference is less clear. Since in December 1938 we caught eight females of over 16·0 g., and among these were five new mice still with prominent teats, it is tempting to think that these or similar late breeders may have been responsible for the batch of small mice characteristic of February 1939 and not found in February 1938. Individual analysis, however, shows that this conspicuous group is as likely to be composed of mice whose spring growth was for some reason delayed, as to be the result of autumn breeding prolonged into December (see also Chart 2, C4, C4X).

(iv) The only other conspicuous difference between the two series of histograms is in March, where among the females of 1938 there is a prominent group of over 20·0 g., absent in 1939. In March 1938 there were three almost unbroken weeks of warm, sunny weather, and towards the end of the month two families were born: one in a cage on 25 March after the mother had been captive for four days, and one in a trap on 26 March. On this we stopped trapping, hoping that we had not already interfered appreciably with the summer population by repeatedly catching and keeping in captivity members of the winter population which might otherwise have started to breed. In 1939, March was a comparatively cold

month, and early April very cold, so that it was not until the beginning of May that we began to get births in the traps and cages (cf. p. 158).

5. Summary

1. Introduction (p. 136).

2. A discussion of the factors which limit the size of the samples caught when trapping *Apodemus* populations for comparison by statistical methods (p. 137).

3. A discussion of the value and reliability of certain conventional measurements, (*a*) in determining the probability that an individual mouse is fully grown, (*b*) in comparing fully grown populations from different localities (pp. 139–44).

The part played by the stomach content (pp. 144–5) and by the reproductive organs (pp. 145–9) in the varying weight of a mouse (p. 141).

4. (*a*) An analysis of monthly weight records of a living population in a single locality (pp. 149–50). These records were made in order to determine the best time of year to trap when comparing populations from different localities; but they also give a picture of the growth of *Apodemus* and how it is influenced by season. The picture is constructed from:

(*b*) Graphs of the weight records of those individual mice caught most frequently in the trapping season 1938–9 (pp. 151–4).

(*c*) Spot diagrams in which the weights, or weight changes, of individual mice are plotted against their weights in a previous month (pp. 154–8 and Charts 1–3).

(*d*) Histograms comparing the weight frequencies of the sample populations caught in each month of the trapping season 1938–9; and a comparison of these histograms with a similar series for 1937–8 when weather conditions were markedly different (pp. 159–61).

Our grateful thanks are due to Miss Enid Harris for her dextrous help in field and laboratory, to Miss Joyce Townend for her skill and care in making the diagrams, to Prof. E. S. Pearson for a critical reading of the manuscript which led to many improvements, and to Dr G. M. Morant and Mrs Karl Pearson for useful suggestions when reading the proofs.

References

Brambell, F. W. R. & Rowlands, I. W. (1936). *Philos. Trans.* B, **226**, 71.

Burt, W. H. (1940). *Misc. Publ. Mus. Zool. Univ. Michigan*, no. 45.

Chitty, D. (1937). *J. Anim. Ecol.* **6**, 36.

Elderton, E. M. & Moul, M. (1928). *Ann. Eugen., Lond.*, **3**, 277.

Elton, C. E., Ford, E. B. & Baker, J. R. (1931). *Proc. Zool. Soc., Lond.*, 657.

Evans, F. C. (1942). *J. Anim. Ecol.* **11**, 182.*

Ruger, H. A. (1933). *Ann. Eugen., Lond.*, **5**, 59.

Ruger, H. A. & Stoessiger, B. (1927). *Ann. Eugen., Lond.*, **2**, 76.

Sumner, F. B. (1926). *J. Mammal.* **7**, 149.

Sumner, F. B. (1927). *J. Mammal.* **8**, 59.

* Received after the present paper had gone to press.

THE CONTROL OF INDUSTRIAL PROCESSES SUBJECT TO TRENDS IN QUALITY

By L. H. C. TIPPETT, *British Cotton Industry Research Association*

In many industrial processes there is a progressive change in the quality of the products with time, which is tolerated up to a certain stage, after which corrective action is taken. An example occurs in the mass production of metal parts by a machining operation to some specified dimension. Random fluctuations in the dimension of articles produced at any one time are caused by such factors as variations in the materials processed and the manipulation of the machine, and superimposed on these is a trend due to tool wear, say in the direction of an increase in the dimension. At some stage, it is decided that too many articles are exceeding the tolerance limit of the dimension and something is done: we shall say that the tool is discarded.

In order to decide when to discard the tool, one procedure is to take samples at regular intervals of time, measure the articles, and calculate the mean dimension for the sample; when that mean reaches a certain value the tool is discarded. But the sample means are subject to sampling errors so that they do not measure exactly the 'true' or population mean of all the articles produced at the given time. Consequently, if tools are discarded at one level of sample means, some will be discarded before the population mean has reached that level and others afterwards; and there will be a frequency distribution of states of wear, as measured by the population mean, at which different tools are discarded. It is the purpose of this paper to determine that frequency distribution and discuss its practical consequences.

The argument will be developed in terms of the above example, but the results may be applied generally to any process having a trend in some characteristic that is estimated at regular intervals by any measure having a sampling error.

The situation is shown diagrammatically in Fig. 1. The line ABC represents the change with time in the population mean, X, of the dimension. X is a characteristic of the tool as set up in the machine, and we may imagine it to be determined, in principle, by making and measuring over a very short interval of time surrounding each instant a very large number of articles, and calculating their mean. In the region AB the trend may be of any form, but in the region BC, which covers the range of the population dimension at which substantially all the tools are discarded, the trend is assumed to be a straight line represented by the equation

$$X = d + ah, \tag{1}$$

where h is time measured in units of the intervals between the taking of successive samples, and d and a are constants. The constant d is chosen to be the level of sample means at which tools are discarded, so that h is measured from the time at which $X = d$. The instants at which samples are taken are

$$(\theta - s),\quad (\theta - s + 1),\quad (\theta - s + 2),\quad \ldots,\quad (\theta - 1),\quad \theta,\quad (\theta + 1),\quad \ldots,$$

where θ is the interval between zero time and the first subsequent sampling instant, and s is an integer defined below.

The sampling distribution of the mean is represented in Fig. 1 for the first three sampling instants, and

$$P(\theta-s),\quad P(\theta-s+1),\quad P(\theta-s+2)$$

are the respective probabilities that the sample means exceed d. For any sampling instant $(\theta-s+v)$ say, $P(\theta-s+v)$ is the probability that a tool, surviving to that instant, will then be discarded. Let $S(\theta-s+v)$ be the probability that any tool will survive to the instant $(\theta-s+v)$ and $D(\theta-s+v)$ the probability that it will be discarded at that instant, so that

$$D(\theta-s+v) = P(\theta-s+v)\,S(\theta-s+v). \tag{2}$$

We shall assume that for a population of tools, θ will have all values between 0 and 1 with equal probability and shall consider the elemental proportion $d\theta$ of those tools for which θ is within the limits $\theta \pm \frac{1}{2}d\theta$. We shall also choose s such that $P(\theta-s) = 0$ to a sufficient degree of accuracy (i.e. that $P(\theta-s) < \epsilon$, where ϵ is a chosen small quantity), and $P(\theta-s+1)$ is

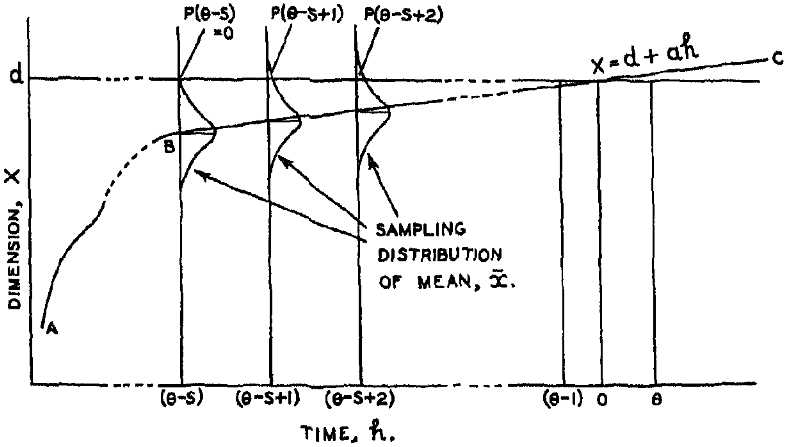


Fig. 1.

the smallest probability of which account need be taken. Then, with the aid of equation (2), we obtain for any tool in the elemental proportion the following probabilities of survival to and discard at the successive sampling instants:

Instant	Probability of survival to instant	Probability of discard at instant
$(\theta-s)$	$S(\theta-s) = 1$	$D(\theta-s) = P(\theta-s)\;S(\theta-s) = 0$
$(\theta-s+1)$	$S(\theta-s+1) = S(\theta-s) - D(\theta-s)$	$D(\theta-s+1) = P(\theta-s+1)\,S(\theta-s+1)$
$(\theta-s+2)$	$S(\theta-s+2) = S(\theta-s+1) - D(\theta-s+1)$	$D(\theta-s+2) = P(\theta-s+2)\,S(\theta-s+2)$
$\vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots \quad \vdots \quad \vdots$

Thus, from the known probabilities $P(\theta-s)$, etc., can be deduced the probabilities $D(\theta-s)$, etc., and these multiplied by $d\theta$ are the elemental probabilities of the times of discard of tools, i.e. they are proportional to the ordinates of the probability distribution of the times of discard.

Probability distributions have been calculated for normal sampling distributions having a constant standard error (S.E.) given by the equation

$$\text{S.E.} = ka, \tag{3}$$

where $k = 1, 2, 5, 10, 20, 33{\cdot}\dot{3}$ and 50 respectively. The probabilities of the time of discard exceeding various values of h are given in Table 1; and a few useful constants, calculated

Table 1. *Probability of time of discard of tool being later than h*

$k=1$		$k=2$		$k=5$		$k=10$		$k=20$		$k=33·3$		$k=50$	
h	Prob.	h	Prob.	h	Prob.	h	Prob.	h	Prob.	h	Prob.	h	Prob.
								−78	1·0000			−196·5	1·0000
						−38	1·0000	−76	0·9999	−129·5	1·0000	−192·5	0·9998
−3·4	1·0000	−6·8	1·0000	−18·5	1·0000	−37	0·9999	−74	0·9997	−126·5	0·9997	−188·5	0·9994
−3·2	0·9999	−6·4	0·9998	−18·0	0·9999	−36	0·9997	−72	0·9994	−123·5	0·9994	−184·5	0·9990
−3·0	0·9997	−6·0	0·9994	−17·5	0·9998	−35	0·9995	−70	0·9990	−120·5	0·9991	−180·5	0·9986
−2·8	0·9993	−5·6	0·9987	−17·0	0·9997	−34	0·9992	−68	0·9985	−117·5	0·9985	−176·5	0·9978
−2·6	0·9986	−5·2	0·9973	−16·5	0·9995	−33	0·9988	−66	0·9977	−114·5	0·9977	−172·5	0·9968
−2·4	0·9974	−4·8	0·9948	−16·0	0·9992	−32	0·9982	−64	0·9965	−111·5	0·9967	−168·5	0·9955
−2·2	0·9953	−4·4	0·9905	−15·5	0·9988	−31	0·9974	−62	0·9948	−108·5	0·9953	−164·5	0·9938
−2·0	0·9917	−4·0	0·9833	−15·0	0·9982	−30	0·9963	−60	0·9926	−105·5	0·9933	−160·5	0·9915
−1·8	0·9859	−3·6	0·9717	−14·5	0·9974	−29	0·9947	−58	0·9893	−102·5	0·9906	−156·5	0·9885
−1·6	0·9770	−3·2	0·9539	−14·0	0·9963	−28	0·9925	−56	0·9848	− 99·5	0·9870	−152·5	0·9845
−1·4	0·9637	−2·8	0·9279	−13·5	0·9948	−27	0·9895	−54	0·9788	− 96·5	0·9821	−148·5	0·9793
−1·2	0·9445	−2·4	0·8909	−13·0	0·9928	−26	0·9855	−52	0·9708	− 93·5	0·9756	−144·5	0·9725
−1·0	0·9178	−2·0	0·8407	−12·5	0·9901	−25	0·9802	−50	0·9603	− 90·5	0·9672	−140·5	0·9641
−0·8	0·8820	−1·6	0·7757	−12·0	0·9865	−24	0·9732	−48	0·9466	− 87·5	0·9561	−136·5	0·9532
−0·6	0·8359	−1·2	0·6954	−11·5	0·9819	−23	0·9640	−46	0·9288	− 84·5	0·9420	−132·5	0·9396
−0·4	0·7787	−0·8	0·6016	−11·0	0·9759	−22	0·9522	−44	0·9062	− 81·5	0·9240	−128·5	0·9226
−0·2	0·7105	−0·4	0·4984	−10·5	0·9681	−21	0·9372	−42	0·8777	− 78·5	0·9015	−124·5	0·9019
0·0	0·6325	0·0	0·3921	−10·0	0·9582	−20	0·9182	−40	0·8425	− 75·5	0·8735	−120·5	0·8764
0·2	0·5473	0·4	0·2904	− 9·5	0·9458	−19	0·8947	−38	0·7999	− 72·5	0·8392	−116·5	0·8457
0·4	0·4583	0·8	0·2006	− 9·0	0·9305	−18	0·8660	−36	0·7492	− 69·5	0·7981	−112·5	0·8090
0·6	0·3698	1·2	0·1279	− 8·5	0·9117	−17	0·8313	−34	0·6902	− 66·5	0·7495	−108·5	0·7660
0·8	0·2862	1·6	0·0745	− 8·0	0·8889	−16	0·7901	−32	0·6234	− 63·5	0·6934	−104·5	0·7163
1·0	0·2116	2·0	0·0392	− 7·5	0·8616	−15	0·7423	−30	0·5501	− 60·5	0·6299	−100·5	0·6600
1·2	0·1487	2·2	0·0273	− 7·0	0·8293	−14	0·6878	−28	0·4722	− 57·5	0·5601	− 96·5	0·5978
1·4	0·0988	2·4	0·0184	− 6·5	0·7918	−13	0·6270	−26	0·3923	− 54·5	0·4855	− 92·5	0·5304
1·6	0·0618	2·6	0·0120	− 6·0	0·7489	−12	0·5608	−24	0·3138	− 51·5	0·4085	− 88·5	0·4596
1·8	0·0362	2·8	0·0076	− 5·5	0·7006	−11	0·4907	−22	0·2401	− 48·5	0·3319	− 84·5	0·3873
2·0	0·0197	3·0	0·0047	− 5·0	0·6473	−10	0·4187	−20	0·1747	− 45·5	0·2589	− 80·5	0·3160
2·1	0·0141	3·2	0·0027	− 4·5	0·5896	− 9	0·3472	−18	0·1200	− 42·5	0·1928	− 76·5	0·2485
2·2	0·0099	3·4	0·0015	− 4·0	0·5284	− 8	0·2787	−16	0·0772	− 39·5	0·1362	− 72·5	0·1873
2·3	0·0068	3·6	0·0008	− 3·5	0·4650	− 7	0·2157	−14	0·0462	− 36·5	0·0905	− 68·5	0·1346
2·4	0·0046	3·8	0·0004	− 3·0	0·4010	− 6	0·1602	−12	0·0255	− 33·5	0·0562	− 64·5	0·0915
2·5	0·0030	4·0	0·0002	− 2·5	0·3381	− 5	0·1137	−10	0·0127	− 30·5	0·0322	− 60·5	0·0585
2·6	0·0019	4·2	0·0001	− 2·0	0·2780	− 4	0·0767	− 8	0·0057	− 27·5	0·0169	− 56·5	0·0349
2·7	0·0012	4·4	0·0000	− 1·5	0·2223	− 3	0·0490	− 6	0·0023	− 24·5	0·0080	− 52·5	0·0192
2·8	0·0007			− 1·0	0·1725	− 2	0·0295	− 4	0·0007	− 21·5	0·0034	− 48·5	0·0097
2·9	0·0004			− 0·5	0·1295	− 1	0·0166	− 2	0·0001	− 18·5	0·0012	− 44·5	0·0044
3·0	0·0002			0·0	0·0938	0	0·0087	0	0·0000	− 15·5	0·0003	− 40·5	0·0017
3·1	0·0001			0·5	0·0654	1	0·0043			− 12·5	0·0000	− 36·5	0·0006
3·2	0·0000			1·0	0·0437	2	0·0020					− 32·5	0·0001
				1·5	0·0279	3	0·0008					− 28·5	0·0000
				2·0	0·0170	4	0·0003						
				2·5	0·0098	5	0·0001						
				3·0	0·0054	6	0·0000						
				3·5	0·0028								
				4·0	0·0014								
				4·5	0·0007								
				5·0	0·0003								
				5·5	0·0001								
				6·0	0·0000								

from a fuller form of the distributions than that given in Table 1, are shown in Table 2. These constants involve the following quantities:

$\bar{h}$, the mean time of discard,

$h_{0\cdot01}$ and $h_{0\cdot05}$ the times of discard exceeded with probabilities of 0·01 and 0·05 respectively, and

$X_{0\cdot01}$ and $X_{0\cdot05}$ the population values of the dimension at the corresponding times, related to the values of h by equation (1).

Values from Table 2 are plotted against k in Figs. 1 and 2 to facilitate interpolation for intermediate values of k†.

Table 2

k	1	2	5	10	20	33·3	50
Mean time of discard, $\bar{h}$	0·27	−0·49	−4·01	−11·61	−29·68	−56·68	−93·16
Time exceeded with probability 0·01:							
$h_{0\cdot01}$*	2·20	2·68	2·50	−0·22	−9·38	−25·33	−48·66
$(h_{0\cdot01}-\bar{h})$	1·93	3·17	6·51	11·39	20·30	31·35	44·50
$(d-X_{0\cdot01})$/S.E.	−2·20	−1·34	−0·50	0·02	0·47	0·76	0·97
Time exceeded with probability 0·05:							
$h_{0\cdot05}$*	1·68	1·86	0·84	−3·04	−14·29	−32·82	−59·21
$(h_{0\cdot05}-\bar{h})$	1·41	2·35	4·85	8·57	15·39	23·86	33·95
$(d-X_{0\cdot05})$/S.E.	−1·68	−0·93	−0·17	0·30	0·71	0·98	1·18

* These are also $(X_{0\cdot01}-d)/a$ and $(X_{0\cdot05}-d)/a$ according to equation (1).

APPLICATIONS

The results of the previous section will usually be applied to control the upper limit of X at which tools are discarded, for it is then that defective articles begin to be produced. The theory does not allow the specification of an absolute upper limit of X beyond which no tools will be used—the assumption of normality allows the theoretical possibility of tools being used up to the stage at which X approaches infinity. It is possible only to specify an upper limit of X that will be exceeded with some probability, and in the following discussion we shall take this probability to be 0·01 considering $X_{0\cdot01}$ to be the limit of X and $h_{0\cdot01}$ as the latest time that tools will be discarded.

First, we may calculate $X_{0\cdot01}$ for any chosen level of discard. The practice in ordinary quality control of setting control limits at 1·96 S.E. above and below the control level‡ suggests, on the face of it, that the level of discard, d, should be at sample means 1·96 S.E. below the upper tolerance limit L given in the specification of this article. Then

$$d = L - 1{\cdot}96\,\text{S.E.},$$

and if we subtract $d - X_{0\cdot01}$ from both sides we have

$$X_{0\cdot01} = L - \text{S.E.}\left\{\frac{d - X_{0\cdot01}}{\text{S.E.}} + 1{\cdot}96\right\}.$$

† The values in Tables 1 and 2 may have errors of 1 or, occasionally, 2 in the last figure.

‡ See Table 10 of pamphlet B.S. 600 R, entitled, *Quality Control Charts*, by B. P. Dudding and W. J. Jennett, published by the British Standards Institution. The inner limits are referred to here.

From Table 2 we see that, if $k = 1$,

$$X_{0\cdot01} = L + 0\cdot24\,\text{S.E.}$$

Thus, more than one tool in a hundred are in use after they are producing articles with a population mean dimension above the tolerance limit, and hence with above 50 % defectives.

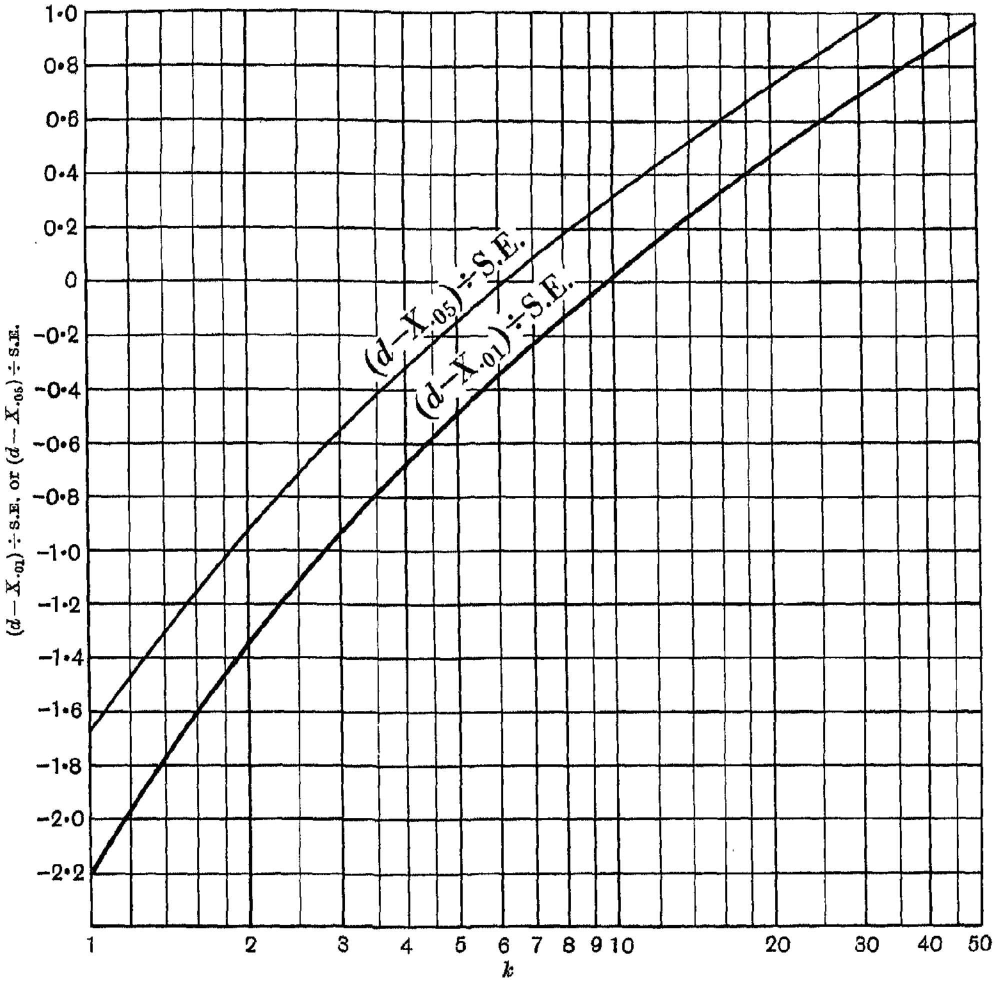


Fig. 2.

The situation is better if the rate of wear is greater compared with the standard error of sample means, so that k is increased. Thus, if $k = 50$,

$$X_{0\cdot01} = L - 2\cdot93\,\text{S.E.}$$

Now, in order to determine the percentage of defectives produced just before the time of discard, let us suppose that the sample size is 16 and that the standard deviation of individual articles made at any one time is σ, so that

$$\text{S.E.} = \tfrac{1}{4}\sigma.$$

Then, for $k = 50$, $$X_{0\cdot01} = L - 0\cdot73\sigma,$$

and if the frequency distribution of individual articles made at any one time is normal, the percentage of defectives at the above value of $X_{0\cdot01}$ is about 22 (Table 9 of B.S. 600 R). This example is sufficient to show that a level of discard related to the tolerance limit in terms of only the standard error of the sample mean is unsatisfactory. The rate of change of X with time and the size of the sample must also be considered.

Usually, it will be more convenient to choose $X_{0\cdot01}$ and use Table 2 to determine the appropriate level of discard. $X_{0\cdot01}$ may be chosen in relation to the engineering limits so that at that value some given percentage of defective articles is produced. The following fictitious example illustrates the procedure.

Example I. Suppose that an article is being produced to a specified diameter with an upper tolerance limit of 6·30 in., that a preliminary investigation has shown that the frequency distribution of articles produced on one machine at one time is substantially normal with a standard deviation of $\sigma = 0\cdot057$ in., that the variability is in control, that tool wear and similar factors cause an average increase in diameter of $a = 0\cdot0005$ in. per hour in the neighbourhood of the time of retooling and resetting, and that eight articles are measured for diameter every hour. At what sample mean diameter, d, must the tool be discarded to ensure that only one tool in a hundred ever produces more than 10 % defective articles through being oversize? What is the average tool life compared with the life that would be attained if all tools could be discarded just when they begin to produce 10 % defective articles?

Table 9 A of B.S. 600 R shows that the value of t, the standardized deviation of the variable distributed normally, corresponding to 10 % defectives is 1·2816, so that the corresponding value of X is $6\cdot30 - 1\cdot2816 \times 0\cdot057 = 6\cdot227$ in.; this is the chosen $X_{0\cdot01}$. For sample of eight, S.E. = 0·0201 in. and $k = 40$; and from Fig. 2 we find that

$$\frac{d - X_{0\cdot01}}{\text{S.E.}} = 0\cdot84,$$

whence $$d = 0\cdot0201 \times 0\cdot84 + 6\cdot227 = 6\cdot244 \text{ in.}$$

Also we find from Fig. 3 that $(h_{0\cdot01} - \bar{h}) = 37$, so that the mean life of the tools is 37 hr. less than it would have been had it been possible to use them all until they produced 10 % defectives.

It should be noted that most of the tools (99 % of them) will never produce as many as 10 % defective articles, and that some (1 % of them) will produce more; for most of the time all tools will produce many fewer than 10 % defectives. Hence, if articles from a number of tools in different stages of wear can be bulked, the average percentage of defectives will be very low. This average can be determined if the (X, h) curve is known throughout the whole of the life of the tool.

The level of discard can be worked out for an $X_{0\cdot01}$ corresponding to any other limit of allowed percentage of defectives, or for a probability level other than 0·01 with which that limit is exceeded. The relevant quantities for a level of 0·05 are given in Table 2 and Figs. 1 and 2, and for other levels they may be obtained from Table 1. The choice of these quantities depends on technical conditions and requirements, and no guiding rules can be given. In practice it will be difficult to make a decision based on measured quantities, and a certain amount of guesswork and arbitrary choice will probably be involved. It would be easy, but laborious, to calculate the mean percentage of defective articles produced by all tools immediately before discard, for various levels of discard, and to use this in choosing the level; but it is doubtful if such a choice would be any easier to make in practice.

The above application considers the approach of articles to the upper tolerance limit only; the lower limit will determine the original setting of the tool in the machine. If the trend is one of decreasing dimension, the signs of all the quantities are reversed and the approach to the lower tolerance limit is controlled. The procedure does nothing towards controlling erratic fluctuations.

Example II. This example concerns the times to melt successive casts of steel in an open-hearth furnace, and the data are taken from *Statistical Methods in Industry*,* Fig. 7 and Table XX. In the figure, the melting times for successive casts are plotted against the cast

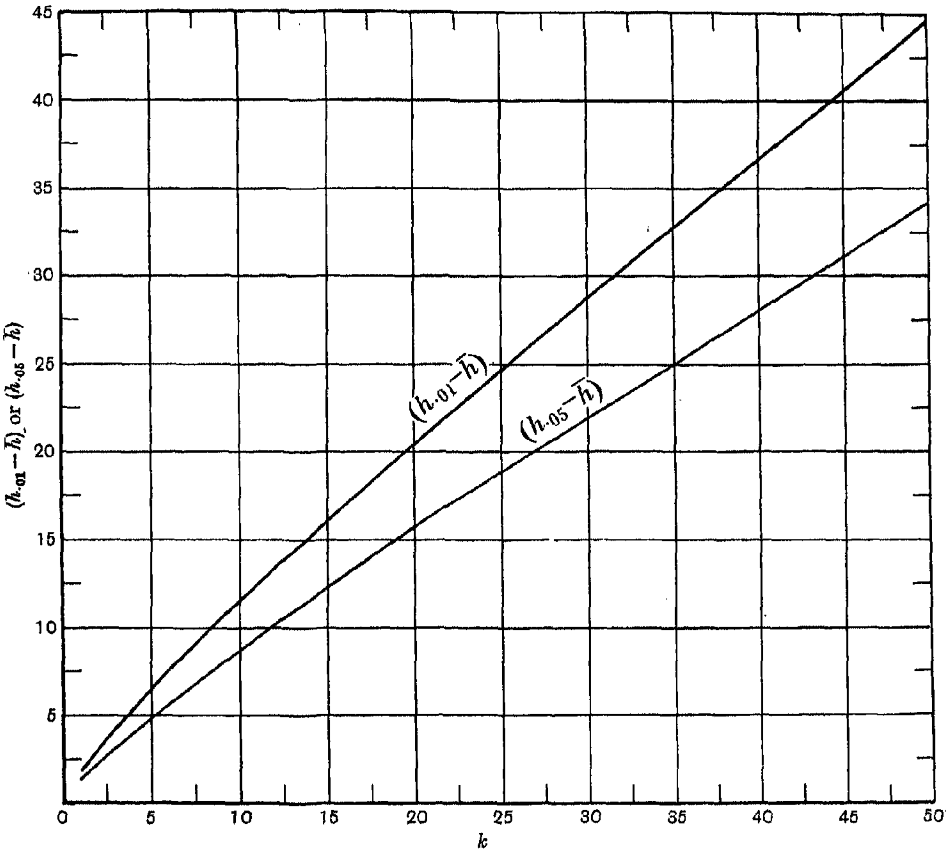


Fig. 3.

number, and for *A* furnace there is a substantial random fluctuation superimposed on a secular movement which appears to be roughly periodic, rising to peaks at intervals of 40–50 casts. A secular variation of this sort could be due to the gradual obstruction of the passages for the gases, which are cleared at intervals. The random variations could be due to such things as variations from cast to cast in the material charged into the furnace. It is not known whether these factors explain the fluctuations for *A* furnace, but for the sake of argument let us suppose that it is so, and let us derive a procedure for deciding when to clean the flues. The data are inadequate to justify this supposition and can provide only very

* Published by the Iron and Steel Industrial Research Council of the British Iron and Steel Federation.

rough estimates of the various quantities; but for the purposes of illustration this does not matter.

Ideally, we would describe the state of the furnace by the time it takes to melt an average cast, undisturbed by random fluctuations; this is our unknown population value, X. The standard deviation of the random fluctuations in the actual times about X is the S.E., and may be roughly estimated from the mean range between consecutive pairs of times. In this way, S.E. is estimated to be 1·17 hr. A rough estimate of a, the rate of increase in melting time, obtained from the figure is 0·16 hr. per cast, so that

$$k = \frac{1 \cdot 17}{0 \cdot 16} = 7$$

(h is the cast number), and from Fig. 2 we find that

$$(d - X_{0\cdot01})/\text{S.E.} = -0\cdot23,$$

whence $$(d - X_{0\cdot01}) = -0\cdot27 \text{ hr.}$$

Suppose it is decided that $X_{0\cdot01}$ shall be 14 hr.; then $d = 13\cdot7$ hr. From Fig. 3 we find that $(h_{0\cdot01} - \bar{h}) = 7\cdot5$ casts, and the average melting time immediately before the furnace is cleaned is $14 - 7\cdot5 \times 0\cdot16 = 12\cdot8$ hr.

There are some conditions to be satisfied before the results of this paper can be applied. The assumption of the constancy of the standard error of sample means involves the assumption that the variability is in control. The mean dimension need not be in control, except for the trend, but it is better that it should be so. Otherwise the variability due to uncontrolled random variations in X must be added to the sampling variations when calculating the S.E. The question of control of X does not arise in circumstances like those envisaged in Example II, where the S.E. is due entirely to random variations in X. Another condition is that a and S.E. are known and that the sampling variation of $\bar{x}$, the sample means, is normal. All these conditions require a preliminary investigation.

Optimum conditions for control

It will be interesting to see whether it is better to take large samples at infrequent intervals or small ones at correspondingly more frequent intervals, the total number of tests over a given length of time remaining constant. For the probability level of 0·01, it is immaterial whether the intervals are frequent or infrequent provided that

$$(h_{0\cdot01} - \bar{h}) = \alpha k^{\frac{2}{3}},$$

where α is a constant. For if the interval between the samples is increased from 1 to r units of time and the sample size is increased to r times its original value, the new standard error is $\text{S.E.}' = \text{S.E.}/\sqrt{r}$, the rate of deterioration in terms of the new units of time is $a' = ra$, the new k becomes $k' = k/r^{\frac{3}{2}}$, and the new $(h_{0\cdot01} - \bar{h})$ becomes $\alpha k'^{\frac{2}{3}} = \alpha k^{\frac{2}{3}}/r$ in the new units, or $\alpha k^{\frac{2}{3}}$ in the original units of time, which is the same as the original value of $(h_{0\cdot01} - \bar{h})$. When the values of $(h_{0\cdot01} - \bar{h})$ in Table 2 are plotted against $k^{\frac{2}{3}}$, they are found to lie on a curve starting near the origin and concave to the $(h_{0\cdot01} - \bar{h})$ axis; i.e. $(h_{0\cdot01} - \bar{h})$ increases more rapidly than in proportion to $k^{\frac{2}{3}}$, and it is better to have a large value of r so as to keep k small. The same is true of $(h_{0\cdot05} - \bar{h})$. Hence, within the limits of this investigation, for a given total number of articles tested in a given total time, it is better, from the point of view of keeping the average tool life as near as possible to the chosen limiting life, to take large samples at relatively infrequent intervals than small samples at frequent intervals.

We may illustrate this by Example I. If samples of thirty-two articles are taken at 4-hourly intervals,

$$a = 4 \times 0{\cdot}0005 = 0{\cdot}0020 \text{ in. per interval},$$

$$\text{S.E.} = \frac{0{\cdot}057}{\sqrt{32}} = 0{\cdot}101 \text{ in.},$$

$$k = 5{\cdot}05,$$

and from Table II,

$$(h_{0\cdot01} - \bar{h}) = 6{\cdot}5 \text{ intervals}$$
$$= 26 \text{ hr. (approx.)}.$$

This may be compared with the 37 hr. for hourly samples of eight articles.

Another method of improving control that suggests itself is to combine each sample mean with the few previous means, using the assumption of a linear trend to obtain an improved estimate of X at the instant when the last sample is taken.

Let the number of means combined in this way be m, and let them be designated

$$\bar{x}_1,\ \bar{x}_2,\ \ldots,\ \bar{x}_u,\ \ldots,\ \bar{x}_m,$$

$\bar{x}_m$ *being the last in the series. Then it is easy to show that if a straight line is fitted to these* values by the method of least squares, the improved estimate of X at the last sampling instant, $\bar{x}'_m$ say, is given by the equation

$$\bar{x}'_m = \frac{2}{m(m+1)} \sum_{u=1}^{u=m} (3u - m - 1)\,\bar{x}_u, \tag{4}$$

and the standard error of this estimate, $\text{S.E.}_{\cdot m}$ say, by the equation

$$\text{S.E.}_{\cdot m} = \sqrt{\frac{2(2m-1)}{m(m+1)}} \times \text{S.E.} \tag{5}$$

Table 3 shows how the precision of the estimate increases with m.

Table 3

m	S.E.$_{\cdot m}$/S.E.	m	S.E.$_{\cdot m}$/S.E.
1	1·00	5	0·77
2	1·00	6	0·72
3	0·91	8	0·65
4	0·84	10	0·59

It is wrong, of course, to combine so many means extending over such a length of time that the departure of the (X, h) curve from linearity for that time is appreciable compared with the sampling errors.

Suppose in Example I the hourly results are combined in this way in fours. Then

$$\text{S.E.}_{\cdot m} = 0{\cdot}84 \times 0{\cdot}0201 = 0{\cdot}0169 \text{ in.},$$

$$k = 33{\cdot}8,$$

and

$$(h_{0\cdot01} - \bar{h}) = 31 \text{ hr. (approx.)}.$$

For this example, the concentration of the thirty-two tests at 4-hourly intervals, giving $(h_{0\cdot01} - \bar{h}) = 26$ hr., remains the best arrangement. Sometimes, however, it is better to test more frequently and combine means by this second method. For example, if with hourly

tests, $k = 31{\cdot}62$, then $(h_{0{\cdot}01} - \bar{h}) = 30$ hr. If ten times as many tests are taken at 10-hourly intervals, k becomes $1{\cdot}0$ and $(h_{0{\cdot}01} - \bar{h}) = 10 \times 2{\cdot}20 = 22$ hr. If the hourly tests are combined in tens, k becomes $0{\cdot}59 \times 31{\cdot}62 = 18{\cdot}6$, and $(h_{0{\cdot}01} - \bar{h}) = 19$ hr. This second method of combination is the only one available in cases like that of Example II.

Summary

1. If samples of a product deteriorating linearly with time are taken at intervals, and action is taken when the sample values reach a certain level, such action will be taken, sometimes before and sometimes after the population value has reached that level.

2. Frequency distributions are given for the times at which action is taken. These distributions depend on the standard error of the sample mean, the size of the sample and the rate of deterioration of the product.

3. A procedure is given for deciding when to take action.

4. For a given number of tests, it is better to take large samples infrequently than small ones frequently. A method is given for combining the results of several consecutive samples.

STUDENTIZATION

OR THE ELIMINATION OF THE STANDARD DEVIATION OF THE PARENT POPULATION FROM THE RANDOM SAMPLE-DISTRIBUTION OF STATISTICS

By H. O. HARTLEY

1. Introduction

In this paper we give a mathematical solution of a problem in statistical distribution, sometimes referred to as 'Studentization'. This process is best described by recalling briefly the properties of two well known examples:

Let $x_1, \ldots, x_n$ denote a sample of n observations drawn from a normal distribution with mean μ and standard deviation σ. Then the difference between sample and population mean, $\bar{x}-\mu$, is normally distributed with zero mean and standard deviation $\sigma/\sqrt{n}$. If we estimate the standard deviation of the parent, σ, by

$$s = \sqrt{[\Sigma(x-\bar{x})^2/(n-1)]},$$

then the ratio

$$t = (\bar{x}-\mu)\sqrt{n}/s$$

is distributed according to Student's t-distribution which does not depend on σ.

Consider now a second, independent sample $X_1, \ldots, X_m$ drawn from the same population. The random sample distribution of its standard deviation

$$s^* = \sqrt{[\Sigma(X-\bar{X})^2/(m-1)]}$$

follows a well-known probability law which, again, involves the value of σ. If, again, we estimate this value of σ by using the standard deviation, s, of the first sample as its estimate, then the distribution of the ratio

$$\sqrt{F} = s^*/s$$

is independent of σ. This distribution is, of course, a transformation of Fisher's z-distribution.

It will be noted that the original statistics, $\bar{x}-\mu$ and s^*, as well as s, are proportional to σ in the sense that if the observations x_i and X_j are measured in units of σ the above statistics are, by this change of scale, transformed into $(\bar{x}-\mu)/\sigma$, s^*/σ and s/σ respectively, and their parental distribution is transformed into the normal distribution with unit standard deviation. Hence the random sample distributions of the ratios

$$t = \frac{\bar{x}-\mu}{\sigma} : \frac{s}{\sqrt{n}\,\sigma} = \frac{(\bar{x}-\mu)}{s}\sqrt{n}, \quad \sqrt{F} = \frac{s^*}{\sigma} : \frac{s}{\sigma} = \frac{s^*}{s}$$

do not involve the value of σ, nor is it necessary to know this value of σ in order to calculate the above ratios. The latter may be called the studentized statistics corresponding to the original statistics $\bar{x}-\mu$ and s^*.

This simple studentization process of dividing a statistic by an independent estimate s of the parental standard deviation σ may be applied to a number of other useful statistics which are also proportional to σ in the above sense. We may mention here the range in a sample, the extreme observation, the extreme value in a sample of random standard deviations, but there are many others.

It is the object of this paper to derive a general formula for the studentized distribution law of such statistics. The accuracy of the formula will be demonstrated with a numerical

example and a new formula for the incomplete Beta function derived in the process. This result has already been used in recent numerical work on this function (Thompson 1941). A second application of the method may be found in a paper by Pearson & Hartley (1943), and it is hoped to publish others.

Finally, computational methods for the systematic tabulation of studentized distribution laws are briefly outlined.

2. Definitions and notation

A problem akin to the present one has been considered in a previous paper (Hartley, 1938), but to fix ideas we shall repeat here certain definitions and results given there.

Let $x_1, \ldots, x_n$ be a sample of n observations drawn from a normal population with standard deviation σ. Consider a general statistic W, calculated from this sample, having the following properties:

(*a*) W is positive.†

(*b*) W is 'proportional to σ' in the sense that if the observations x_i are measured in units of σ the statistic W is thereby transformed into W/σ.

Let $f(W)$ be the distribution function of W for the case where σ is known to be unity, and let $P(W)$ be the corresponding probability integral, i.e. $P(W) = \int_0^W f(W)\,dW$. Also let s be an independent estimate of the standard deviation based on n degrees of freedom. Finally, let‡

$$S = \sqrt{n}\,s = \sqrt{(\text{Sum of squares})} \quad \text{and} \quad r = W/S.$$

Further denote by $f_n(r)$ the random sample distribution of r and by $p_n(R)$ its probability integral, i.e.

$$p_n(R) = \int_0^R f_n(r)\,dr.$$

Using this notation we proved (see Hartley, 1938) that

$$p_n(R) = \Gamma(\tfrac{1}{2}n)^{-1}\,2^{-\frac{1}{2}n+1}\int_0^\infty S^{n-1}e^{-\frac{1}{2}S^2}P(SR)\,dS. \tag{1}$$

This equation gives $p_n(R)$, the 'studentized integral', formally in terms of the '∞-integral' $P(W)$. We also proved the recurrence formula

$$p_{n-2}(R) = (n-2)\,R^{-(n-2)}\int_0^R p_n(r)\,r^{n-3}\,dr, \tag{2}$$

with the help of which the degrees of freedom of $p_n(R)$ may be reduced by 2.

3. Properties of the recurrence formula

The series of studentized integrals $p_n(R)$ is uniquely determined by the recurrence formula (2) and the relations

$$\lim_{n\to\infty} p_n\left(\frac{W}{\sqrt{n}}\right) = P(W) \quad \text{(uniformly in } W\text{)}, \quad p_n \text{ bounded}, \tag{3}$$

which follow from an inequality derived previously (see Hartley, 1938, (10)) and from (2). To prove this suppose there were two series of piecewise continuous functions $p_n(R)$ and

† It will be seen that this condition is not essential for our results. Also, if the distribution function of W is symmetrical about 0, condition (*a*) will be satisfied for the statistic $|W|$ which may be used in place of W.

‡ It was necessary to alter the notation used previously (Hartley, 1938) writing S in place of ρ and r in place of q, in order to fall in with the notation used by other writers (e.g. Newman, 1939) who reserve q for the ratio W/s.

$\pi_n(R)$ both satisfying (2) and (3), and let us denote by Δ_n the maximum of $|p_n(R)-\pi_n(R)|$ for $0 \leqslant R < \infty$. From (3) we obtain

$$\lim_{n\to\infty} \Delta_n = 0. \tag{4}$$

On the other hand, we find from equation (2) that

$$p_{n-2}(R)-\pi_{n-2}(R) = (n-2)\,R^{-(n-2)}\int_0^R [p_n(r)-\pi_n(r)]\,r^{n-3}\,dr,$$

from which we obtain $$\Delta_{n-2} \leqslant (n-2)\,R^{-(n-2)}\Delta_n \frac{R^{n-2}}{n-2} = \Delta_n,$$

i.e. $$\Delta_{n-2} \leqslant \Delta_n. \tag{5}$$

From (4) and (5) it follows that

$$\Delta_n = 0$$

for all n.

The recurrence formula (2) may be used to produce by quadrature $p_{n-2}(R)$ from $p_n(R)$. The difficulty is that none of the integrals $p_n(R)$ is known so that we have no starting point. All we know is that

$$\lim_{n\to\infty} p_n\left(\frac{W}{\sqrt{n}}\right) = P(W).$$

This suggests that a first and main step of the process will consist in bridging the gap between $n=\infty$ (i.e. $P(W)$) and moderate values of n (i.e. $p_n(R)$), say, covering the range between 10 and 50. This step will be accomplished with the help of the theory of partial differential equations in the next section.

4. The partial differential equation of $p_n(R)$

We may write formula (2) in the form

$$R^n p_n(R) = n\int_0^R p_{n+2}(r)\,r^{n-1}\,dr. \tag{6}$$

Differentiating with regard to R we obtain

$$R\frac{dp_n}{dR} - n[p_{n+2}(R)-p_n(R)] = 0. \tag{7}$$

We now regard the degrees of freedom n as a continuous second variable capable of attaining positive values and try to find a function of two variables $p(n,R)$ for which the relations

$$R\frac{\partial p(n,R)}{\partial R} - n[p(n+2,R)-p(n,R)] = 0 \tag{8}$$

and $$\lim_{n\to\infty} p\left(n, \frac{W}{\sqrt{n}}\right) = P(W) \tag{9}$$

hold. Again this function $p(n,R)$ is uniquely determined by the relations (8) and (9). This is easily seen from integrating (8) which yields

$$R^n p(n,R) = n\int_0^R p(n+2,r)\,r^{n-1}\,dr + g(n).$$

Putting $R=0$ we see that $g(n)\equiv 0$, i.e. that equation (6) holds. The remainder follows by the argument given in § 3.

To convert equation (8) into a partial differential equation we expand the finite difference in a Taylor series and write

$$R\frac{\partial p}{\partial R}-n\left[2\frac{\partial p}{\partial n}+2\frac{\partial^2 p}{\partial n^2}+\frac{4}{3}\frac{\partial^3 p}{\partial n^3}+\frac{2}{3}\frac{\partial^4 \overline{p}}{\partial n^4}\right]=0, \tag{10}$$

where p is taken at arguments R and n whilst $\overline{p}$ is taken at R and some mean value between n and $n+2$.

We now introduce as new variables

$$y=\log R \quad \text{and} \quad x=\log n.$$

The partial differential equation is thereby transformed into

$$\frac{\partial p}{\partial y}-2\frac{\partial p}{\partial x}=2e^{-x}\left(\frac{\partial^2 p}{\partial x^2}-\frac{\partial p}{\partial x}\right)+\tfrac{4}{3}e^{-2x}\left(\frac{\partial^3 p}{\partial x^3}-3\frac{\partial^2 p}{\partial x^2}+2\frac{\partial p}{\partial x}\right)$$
$$+\tfrac{2}{3}e^{-3x}\left(\frac{\partial^4 \overline{p}}{\partial x^4}-6\frac{\partial^3 \overline{p}}{\partial x^3}+11\frac{\partial^2 \overline{p}}{\partial x^2}-6\frac{\partial \overline{p}}{\partial x}\right), \tag{11}$$

where the arguments of p are y and x and those of $\overline{p}$ are y and some mean value x between x and $\log(e^x+2)$.

5. The solution of the partial differential equation by iteration

The partial differential equation (11) cannot, of course, be solved by standard methods. However, it is possible to find an approximate solution for large and moderate values of x with the help of an iteration process. For $x\to\infty$ the equation (11) tends to the limiting form

$$\frac{\partial p}{\partial y}-2\frac{\partial p}{\partial x}=0, \tag{12}$$

the general solution of which is given by

$$p_0(x,y)=\phi(y+\tfrac{1}{2}x),$$

where $\phi(\lambda)$ is an arbitrary differentiable function of λ. In order to satisfy condition (9) we have to put

$$\phi(\lambda)=P(e^{\lambda}), \tag{13}$$

so that we obtain for the solution $p_0(x,y)$

$$p_0(x,y)=P(e^{y+\frac{1}{2}x}). \tag{14}$$

The first approximation $p_0(x,y)$ to the studentized integral† is therefore a function of $(y+\frac{1}{2}x)$ only i.e., expressed in terms of R and n, p_0 depends on $R\sqrt{n}$ only. This means that for large n the studentized integral is, to a first approximation, independent of n.

In order to work out closer approximations we have to solve the non-homogeneous partial differential equation

$$\frac{\partial p}{\partial y}-2\frac{\partial p}{\partial x}=\psi(y,x). \tag{15}$$

A partial solution of this equation can be found with the help of the theory of characteristics and is given by

$$p(x,y)=\int_0^y \psi(\rho,-2\rho+2y+x)\,d\rho, \tag{16}$$

† We shall confine ourselves here to working out the first three steps of the iteration process. It would lead us too far afield if we would give here the proof for its convergence. The check on the accuracy of the approximation is discussed in the last section.

as is easily verified by differentiation. The general integral of equation (15) is therefore given by

$$p(x,y) = \int_{y_0}^{y} \psi(\rho, -2\rho+2y+x)\,d\rho + \phi(y+\tfrac{1}{2}x), \tag{17}$$

where ϕ, y_0 are arbitrary and ρ is a variable of integration. If the function $\psi(y,x)$ is of the special form

$$\psi(y,x) = \sum_{\nu=1}^{N} e^{-x\nu}\,\psi_\nu(y+\tfrac{1}{2}x), \tag{18}$$

and if we put $y_0 = -\infty$ we find by substituting (18) in (17) and by integrating each term

$$p(x,y) = \phi(y+\tfrac{1}{2}x) + \sum_{\nu=1}^{N} \frac{e^{-x\nu}}{2\nu}\,\psi_\nu(\tfrac{1}{2}x+y). \tag{19}$$

The second approximation $(p_1(x,y))$ to the solution of (11) is now obtained by solving

$$\frac{\partial p_1}{\partial y} - 2\frac{\partial p_1}{\partial x} = 2e^{-x}\left(\frac{\partial^2 p_0}{\partial x^2} - \frac{\partial p_0}{\partial x}\right), \tag{20}$$

using $p_0(x,y) = \phi(y+\tfrac{1}{2}x)$. From (15), (18) and (19) we obtain

$$p_1(x,y) = \phi(y+\tfrac{1}{2}x) + e^{-x}(\tfrac{1}{4}\phi'' - \tfrac{1}{2}\phi'), \tag{21}$$

where we denote by $'$, $''$, $'''$, etc., the order of the derivative of the function ϕ with regard to its argument. The function $p_1(x,y)$ is a solution of the equation (11) if all terms involving e^{-2x}, e^{-3x}, etc., are ignored.

The third approximation $p_2(x,y)$ is now obtained as a solution of the equation

$$\frac{\partial p_2}{\partial y} - 2\frac{\partial p_2}{\partial x} = 2e^{-x}\left(\frac{\partial^2 p_1}{\partial x^2} - \frac{\partial p_1}{\partial x}\right) + \tfrac{4}{3}e^{-2x}\left(\frac{\partial^3 p_1}{\partial x^3} - 3\frac{\partial^2 p_1}{\partial x^2} + 2\frac{\partial p_1}{\partial x}\right), \tag{22}$$

where, in the right-hand side, we may ignore all terms involving e^{-3x}, e^{-4x}, etc. From equations (18) and (22) we obtain for the solution (19)

$$p_2(x,y) = \phi + e^{-x}(\tfrac{1}{4}\phi'' - \tfrac{1}{2}\phi') + e^{-2x}(\tfrac{1}{32}\phi^{\text{iv}} - \tfrac{5}{24}\phi''' + \tfrac{3}{8}\phi'' - \tfrac{1}{6}\phi'), \tag{23}$$

which is a solution of (11) if all terms involving e^{-3x}, e^{-4x}, etc. (i.e. $1/n^3$, $1/n^4$, etc.), are ignored. For samples of the order 20 or larger this formula will in most cases give quite sufficient accuracy. For smaller samples or high accuracy the term involving e^{-3x} will be required. We find by a further iteration

$$p_3(x,y) = p_2(x,y) + \frac{e^{-3x}}{12}(\tfrac{1}{32}\phi^{\text{vi}} - \tfrac{21}{48}\phi^{\text{v}} + \tfrac{17}{8}\phi^{\text{iv}} - \tfrac{11}{6}\phi''' + 3\phi''). \tag{24}$$

Let us write formula (23) in terms of the statistic

$$R\sqrt{n} = e^{(y+\frac{1}{2}x)}.$$

We have the following relations between the function $\phi(\lambda)$ and the ∞-integral $P(W)$ $(\lambda = y+\tfrac{1}{2}x)$

$$\begin{aligned}
\phi &= P(e^{\lambda}),\\
\phi' &= P'e^{\lambda},\\
\phi'' &= P''e^{2\lambda} + P'e^{\lambda},\\
\phi''' &= P'''e^{3\lambda} + 3P''e^{2\lambda} + P'e^{\lambda},\\
\phi^{\text{iv}} &= P^{\text{iv}}e^{4\lambda} + 6P'''e^{3\lambda} + 7P''e^{2\lambda} + P'e^{\lambda}.
\end{aligned}$$

Substituting in equation (23) we obtain the chance $p_n(R)$ for the statistic $r = W/s\sqrt{n}$ to fall below R which is identical with the chance for the studentized statistic W/s to fall below $R\sqrt{n} = Q$ (say). We obtain from the second approximation, i.e. from formula (23)

$$p_n(R) = P(Q) + \frac{1}{4n}(Q^2P'' - QP') + \frac{1}{16n^2}(\tfrac{1}{2}Q^4P^{\mathrm{iv}} - \tfrac{1}{3}Q^3P''' - \tfrac{1}{2}Q^2P'' + \tfrac{1}{2}QP'), \tag{25}$$

where the argument of P and its derivatives is Q, since

$$\phi(y + \tfrac{1}{2}x) = P(e^{y+\frac{1}{2}x}) = P(R\sqrt{n}) = P(Q).$$

With equations (25) and (24) we have reached the solution of the studentization which is of sufficient accuracy for most practical purposes. In the following sections we shall show with an example how these equations may be used in practical instances. We shall see that the formulae lend themselves very well to the computation of $p_n(R)$. For a fixed value of Q $p_n(Q/\sqrt{n})$ is a quadratic in $1/n$ if equation (25) is used. The formula is of a fair accuracy even for small values of n. However, if higher accuracy is required it is preferable to include the term involving $1/n^3$ which is given by equation (24).

6. A new formula for the incomplete B-function†

As is well known the incomplete B-function, the probability integral of Fisher's z and of Snedecor's F, are all transformations of the same probability law which is the studentized integral of a sample standard deviation s_1. If s_1^2 and s_2^2 are both estimates of σ^2 based on n_1 and n_2 degrees of freedom respectively, then the probability that the ratio s_1^2/s_2^2 falls below F is given by

$$I_{x_0}(\tfrac{1}{2}n_1, \tfrac{1}{2}n_2),$$

where I_{x_0} is the incomplete B-function and

$$x_0 = \frac{n_1 F}{n_2 + n_1 F}. \tag{26}$$

Whilst for large p and q there are a number of useful methods facilitating the approximate evaluation of $I_x(p, q)$ (see e.g. Soper, 1921 and Wishart, 1927) there appears to be a lack of suitable methods if p is small or moderate and q is large. This is a case where our method of studentization may be usefully employed.

The distribution of s_1 for n_1 degrees of freedom is given by

$$f(s_1) = \{\Gamma(\tfrac{1}{2}n_1)\}^{-1}\, 2^{-\frac{1}{2}n_1+1} n_1^{n_1} s_1^{n_1-1} e^{-\frac{1}{2}s_1^2 n_1}. \tag{27}$$

All we have to do therefore is to substitute in equation (25) $f(s_1)$ for P' and the derivatives of $f(s_1)$ for P'', P''' and P^{iv} respectively. The argument Q is to be replaced by the square root of the variance ratio F and n by n_2, whilst P itself is the probability integral of $\sqrt{(\chi^2/n_1)}$ for n_1 degrees of freedom.

We find without difficulty

$$p_{n_2}(\sqrt{(Fn_2)}) = P^*(\sqrt{F}) + \frac{1}{4n_2}[Ff' - \sqrt{F}f] + \frac{1}{16n_2^2}(\tfrac{1}{2}F^2f''' - \tfrac{1}{3}F^{\frac{3}{2}}f'' - \tfrac{1}{2}Ff' + \tfrac{1}{2}\sqrt{F}f), \tag{28}$$

where f is given by (27) and the functions f', f'' and f''' are its derivatives whilst the argument s_1 is replaced by $\sqrt{F}$. The integral P^* is the probability integral of s_1 or that of $\sqrt{(\chi^2/n_1)}$.

† Since this paper was written my attention has been drawn to a paper by A. G. Campbell (1923) in which a similar expression has been derived. Campbell's formula is mentioned and used in the recent book by Simon (1941).

It is most convenient to adopt the notation which has been introduced by Karl Pearson in his work on the incomplete B-function. The integral p_n will then become an approximation to the function $I_{x_0}(p,q)$ where $2p = n_1$, $2q = n_2$ and x_0 is given by (26). If, further,

$$\omega = \tfrac{1}{2}Fn_1 = \frac{qx_0}{1-x_0}, \tag{29}$$

we obtain from equations (27) and (28) the new formula for the function $I_{x_0}(p,q)$, viz.

$$\begin{aligned} I_{x_0}(p,q) \equiv {} & P^*(\sqrt{F}) + \tfrac{1}{2}e^{-\omega}\omega^p q^{-1}\Gamma(p)^{-1}(p-1-\omega) \\ & + \tfrac{1}{4}e^{-\omega}\omega^p q^{-2}\Gamma(p)^{-1}\{\tfrac{1}{2}p^3 - \tfrac{5}{3}p^2 + \tfrac{3}{2}p - \tfrac{1}{3} - \omega(\tfrac{3}{2}p^2 - \tfrac{11}{6}p + \tfrac{1}{3}) + \omega^2(\tfrac{3}{2}p - \tfrac{1}{6}) - \tfrac{1}{2}\omega^3\}, \end{aligned} \tag{30}$$

where $P^*(\sqrt{F})$ is the probability integral of $\sqrt{(\chi^2/n_1)}$ for $2p$ degrees of freedom.

The formula is of considerable help for large values of q and small or moderate values of p. If p and ω are kept constant $I_{x_0}(p,q)$ is seen to be a quadratic in $1/q$.

To give some idea of the remarkable accuracy of the formula we consider the example $p = 1$, $\omega = 5\cdot555$ and $q = 5(1)\,10(2\cdot5)\,25(5)\,50$. A comparison of accurate value and approximation is given in the table below. The method has been used to calculate values of the incomplete B-function for $q = 120$ and small values of p. These were required for the new table of percentage levels of $I_{x_0}(p,q)$ prepared by *Biometrika* (Thompson, 1941).

$2q$	$I_x(1,q)$ Exact	$I_x(1,q)$ Approx.	$2q$	$I_x(1,q)$ Exact	$I_x(1,q)$ Approx.
100	0·994 8462	0·994 8452	30	0·991 140	0·991 093
90	6916	6901	25	0·989 922	0·989 830
80	4952	4930	20	0·987 95	0·987 77
70	2374	2342	18	0·986 79	0·986 55
60	0·993 8847	0·993 8795	16	0·985 29	0·984 94
50	374	365	14	0·983 26	0·982 73
45	022	010	12	0·980 41	0·979 54
40	0·992 572	0·992 553	10	0·976 26	0·974 63
35	0·991 973	0·991 943			

7. SOME REMARKS ON THE TABULATION OF 'STUDENTIZED INTEGRALS'

Formulae (24) and (25) are really the first three or four terms of an expansion of the studentized integral in powers of $1/n$, and it is necessary to know something about the accuracy of this approximation. The neglected remainder of the expansion depends on the derivatives of the ∞-integral $P(W)$ and it is difficult to reach a general formula which may be used as a convenient gauge for the estimation of its magnitude. Nevertheless, it will be possible to control, numerically, the accuracy of formula (25). Thus, if it is desired to tabulate the studentized integral to a certain decimal accuracy (say 3 or 4 or 5 decimal accuracy) one would proceed on the following lines:

(i) From the magnitudes of the coefficients of n^0, n^{-1}, n^{-2} and n^{-3} in equations (25) and (24) *estimate roughly* the smallest value of n (n' say) for which equation (25) represents $p_n(R)$ sufficiently accurately for large values of R.

(ii) For two or three values of n in the neighbourhood of the above value n' and paired with two or three values of R, calculate the exact studentized integral $p_n(R)$ from equation (1) by numerical quadrature and compare the result with the value of $p_n(R)$ given

by (25). This comparison should *fix* the smallest value of n (n_0 say) for which the approximate formula (25) yields the desired decimal accuracy.

(iii) For $n \geqslant n_0+2$ use formula (25) for the tabulation of $p_n(R)$. For $n < n_0+2$ use the recurrence formula (2) (starting from n_0+2 and n_0+1 respectively) to produce in turn the exact $p_n(R)$ by numerical quadrature† reducing the degrees of freedom by 2 in each step of the recurrence. The first step of this process should produce the studentized integral $p_n(R)$ and this should agree identically with the one calculated from formula (25).

Finally, we should mention a simple transformation of the recurrence formula (2) which makes it possible to carry out the quadrature (iii) with the help of a mechanical integrator known as the 'Differential Analyser'. A good description of this machine is given by Hartree (1935). We first transform the R-scale by introducing $\rho = R^{-2}$. We then replace the function $p_n(R) = p_n(\rho^{-\frac{1}{2}})$ by

$$\pi_n(\rho) = [1 - p_n(\rho^{-\frac{1}{2}})]\,\rho^{-\frac{1}{2}n}. \tag{31}$$

The recurrence formula (2) is thereby transformed into

$$\pi_n(\mathrm{P}) = \tfrac{1}{2}n\int_{\mathrm{P}}^{\infty} \pi_{n+2}(\rho)\,d\rho. \tag{32}$$

Thus $\pi_n(\rho)$ is seen to be a multiple of the integral of $\pi_{n+2}(\rho)$ and the recurrence process is reduced to ordinary integration.

This process is easily carried out on the Differential Analyser. For instance, with a capacity of six 'integrator units', six 'output-tables' and one 'input table', six integrations may be carried out simultaneously and π_{n-2}, π_{n-4}, π_{n-6}, π_{n-8}, π_{n-10} and π_{n-12} produced from $\pi_n(\rho)$ in one 'run'.

REFERENCES

Campbell, G. A. (1923). *Bell. Syst. Tech. J.* January, 1923.
Hartley, H. O. (1938). *J.R. Statist. Soc. Suppl.* **5**, 80.
Hartree, D. R. (1935) *Nature, Lond.*, **135**, 940.
Newman, D. (1939). *Biometrika*, **31**, 20.
Pearson, E. S. & Hartley, H. O. (1943). *Biometrika*, **33**, 89.
Simon, L. E. (1941). *An Engineer's Manual of Statistical Methods.* New York: John Wiley and Sons, Inc.
Soper, H. E. (1921). *Tracts for Computers*, No. vii. Camb. Univ. Press.
Thompson, C. M. (1941). *Biometrika*, **32**, 151.
Wishart, J. (1927). *Biometrika*, **19**, 1.

† Note that with the recurrence process (2) only one numerical quadrature is required to produce $p_{n-2}(R)$ for all values of R, whilst the exact formula (1) necessitates a separate quadrature over the range 0 to ∞ for every single value of R.

MISCELLANEA

Note on the use of the tables of percentage points of the incomplete beta function to calculate small sample confidence intervals for a binomial p

By HENRY SCHEFFÉ, *Princeton University*

One of many once tedious statistical computations which may now be made simply and directly by the use of Miss Thompson's recent tables (1941) is the calculation of confidence intervals for a binomial p when the sample size is small, so that one hesitates to use the normal approximation to the binomial distribution. The problem was treated by C. J. Clopper & E. S. Pearson (1934). In their article they included charts yielding confidence intervals for $\epsilon = 0{\cdot}01$ and $0{\cdot}05$, where $1-\epsilon$ denotes the confidence coefficient. With the new tables these cases as well as $\epsilon = 0{\cdot}02$, $0{\cdot}1$, $0{\cdot}2$, $0{\cdot}5$ are easily handled.

Let x be the number of 'successes' observed in a sample of n trials on a binomial population for which $E(x/n) = p$. Denote by $p_1(x) \leqq p \leqq p_2(x)$ a confidence interval for p with confidence coefficient $1-\epsilon$. From the work of Clopper & Pearson we find that $p_2(x)$ is determined by the equation

$$\sum_{j=0}^{x} {}_nC_j p_2^j (1-p_2)^{n-j} = \tfrac{1}{2}\epsilon \quad (x<n), \tag{1}$$

while $p_2(n) = 1$, and that $p_1(x)$ may then be found from

$$p_1(x) = 1 - p_2(n-x). \tag{2}$$

Karl Pearson (1924) showed how the left member of (1) may be evaluated in terms of the incomplete Beta function,

$$\sum_{j=0}^{x} {}_nC_j p^j (1-p)^{n-j} = I_{1-p}(n-x, x+1) \quad (x<n). \tag{3}$$

From (1), (2), (3) we deduce the following rule for calculating the confidence limits $p_1(x)$ and $p_2(x)$ from Miss Thompson's tables: Enter the table for the $100(\frac{1}{2}\epsilon)$ percentage point with $\nu_1 = 2(n-x+1)$, $\nu_2 = 2x$ to read $p_1(x)$ directly; in the same table subtract from unity the entry for $\nu_1 = 2(x+1)$, $\nu_2 = 2(n-x)$ to get $p_2(x)$. The exceptions to this rule occur for $p_1(0)$ and $p_2(n)$, to which we assign the values 0 and 1, respectively. If the percentage point for the desired ν_1, ν_2 is not tabulated, Hartley's *Methods of Interpolation* preceding the tables will yield it fairly quickly, especially since only two decimals will be wanted ordinarily.

REFERENCES

Clopper, C. J. & Pearson, E. S. (1934). *Biometrika*, **26**, 404–13.
Pearson, Karl (1924). *Biometrika*, **16**, 202–3.
Thompson, Catherine M. (1941). *Biometrika*, **32**, 151–81.

BOOKS RECEIVED FOR REVIEW

The Genetics of the Mouse. By HANS GRÜNEBERG. Cambridge: The University Press. 1943. Price 30*s*.

The Advanced Theory of Statistics. Vol. I. By MAURICE G. KENDALL. London: Charles Griffin and Co. Ltd. 1943. Price 42*s*.

The Statistical Study of Literary Vocabulary. By G. UDNY YULE. Cambridge: The University Press. 1944. Price 25*s*.

The Year Book of Labour Statistics. Seventh year of issue. 1942. Montreal: International Labour Office. 1943. Price 8*s*.

(1) *Table of Circular and Hyperbolic Tangents and Cotangents for Radian Arguments*. (2) *Table of the Bessel Functions $J_0(z)$ and $J_1(z)$ for Complex Arguments*. (3) *Tables of Lagrangian Interpolation Coefficients*. (4) *Table of Reciprocals of Integers from* 100,000 *through* 200,009. Tables prepared by the Mathematical Tables Project, Works Projects Administration of the Federal Works Agency, conducted under the sponsorship of the National Bureau of Standards and edited by LYMAN J. BRIGGS and ARNOLD N. LOWAN. New York: Columbia University Press. 1943–44. Price: (1)–(3) \$5.00 each volume, (4) \$4.00.

Vol. XXXIII. Part III

November, 1945

Dept. of Psychological Foundations
N.C.E.R.T.
RG

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

ASSISTED BY

EGON S. PEARSON

Reprinted by offset-litho

ISSUED BY THE BIOMETRIC LABORATORY
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

VOLUME XXXIII, PART III NOVEMBER 1945

ON THE USE OF MATRICES IN CERTAIN POPULATION MATHEMATICS

BY P. H. LESLIE, *Bureau of Animal Population, Oxford University*

CONTENTS

1. INTRODUCTION

If we are given the age distribution of a population on a certain date, we may require to know the age distribution of the survivors and descendants of the original population at successive intervals of time, supposing that these individuals are subject to some given age-specific rates of fertility and mortality. In order to simplify the problem as much as possible, it will be assumed that the age-specific rates remain constant over a period of time, and the female population alone will be considered. The initial age distribution may be entirely arbitrary; thus, for instance, it might consist of a group of females confined to only one of the age classes.

The method of computing the female population in one unit's time, given any arbitrary age distribution at time t, may be expressed in the form of $m+1$ linear equations, where m to $m+1$ is the last age group considered in the complete life table distribution, and when the same unit of age is adopted as that of time. If

n_{xt} = the number of females alive in the age group x to $x+1$ at time t,

P_x = the probability that a female aged x to $x+1$ at time t will be alive in the age group $x+1$ to $x+2$ at time $t+1$,

F_x = the number of daughters born in the interval t to $t+1$ per female alive aged x to $x+1$ at time t, who will be alive in the age group 0–1 at time $t+1$,

then, working from an origin of time, the age distribution at the end of one unit's interval will be given by

$$\begin{aligned}
\sum_{x=0}^{m} F_x n_{x0} &= n_{01} \\
P_0 n_{00} &= n_{11} \\
P_1 n_{10} &= n_{21} \\
P_2 n_{20} &= n_{31} \\
\vdots \quad & \quad \vdots \\
P_{m-1} n_{m-1,0} &= n_{m1}
\end{aligned}$$

or, employing matrix notation, $Mn_0 = n_1$, where n_0 and n_1 are column vectors giving the age distribution at $t = 0$ and 1 respectively, and the matrix

$$M = \begin{bmatrix} F_0 & F_1 & F_2 & \cdots & . & F_k & F_{k+1} & \cdots & F_{m-1} & F_m \\ P_0 & . & . & \cdots & . & . & . & \cdots & . & . \\ . & P_1 & . & \cdots & . & . & . & \cdots & . & . \\ . & . & P_2 & \cdots & . & . & . & \cdots & . & . \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ . & . & . & \cdots & P_{k-1} & . & . & \cdots & . & . \\ . & . & . & \cdots & . & P_k & . & \cdots & . & . \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ . & . & . & \cdots & . & . & . & \cdots & P_{m-1} & . \end{bmatrix} \quad 0 < P_x < 1;\ F_x \geqq 0.$$

This matrix is square and consists of $m+1$ rows and $m+1$ columns. All the elements are zero, except those in the first row and in the subdiagonal immediately below the principal diagonal. The P_x figures all lie between 0 and 1, while the F_x figures are by definition necessarily positive quantities. Some of the latter, however, may be zero, their number and position depending on the reproductive biology of the species we happen to be considering in any particular case, and on the relative span of the pre- and post-reproductive ages. If $F_m = 0$, the matrix M is singular, since the determinant $|M| = 0$.

Since $Mn_0 = n_1$, and $Mn_1 = M^2_{n_0} = n_2$, etc., the age distribution at time t may be found by pre-multiplying the column vector $\{n_{00}\, n_{10}\, n_{20} \ldots n_{m0}\}$, i.e. the age distribution at $t = 0$, by the matrix M^t. Moreover, it will be seen that with the help of the jth column of M^t the age distribution and number of the survivors and descendants of the $n_{j-1,0}$ individuals, who were alive at $t = 0$, can readily be calculated. Thus, $n_{j-1,0}$ times the sum of the elements in the jth column of M^t gives the number of living individuals contributed to the total population at time t by this particular age group.

2. Derivation of the matrix elements

The basic data, from which the numerical elements of this matrix may be derived, are given usually in the form of a life table and a table of age specific fertility rates. To take the P_x figures first; if at $t = 0$ there are n_{x0} females alive in the age group x to $x+1$, the survivors of these will form the $x+1$ to $x+2$ age group in one unit's time, and thus $P_x n_{x0} = n_{x+1,1}$. Then it is usually assumed (e.g. Charles, 1938, p. 79; Glass, 1940, p. 464) that

$$P_x = \frac{L_{x+1}}{L_x},$$

where

$$L_x = \int_x^{x+1} l_x \, dx,$$

or the number alive in the age group x to $x+1$ in the stationary or life table age distribution. This method of computing the survivors in one unit's time would be exact if the distribution of those alive within a particular age group was the same as in the life-table distribution.

The F_x figures are more troublesome, and in the numerical example which will be given later they were obtained from the basic maternal frequency figures (m_x = the number of live daughters born per unit of time to a female aged x to $x+1$) by an argument which ran as follows. Consider the n_{x0} females alive at $t = 0$ in the age group x to $x+1$, and let us sup-

pose that they are concentrated at the midpoint of the group, $x+\frac{1}{2}$. During the interval of time 0–1 some of these individuals are dying off, and at $t = 1$ the $n_{x+1,1}$ survivors can be regarded as concentrated at the age $x+1\frac{1}{2}$. Although these deaths are taking place continuously, we may assume them all to occur around $t = \frac{1}{2}$, so that at this latter time the number of females alive in the age group we are considering changes abruptly from n_{x0} to $n_{x+1,1} = P_x n_{x0}$. Then during the time interval $0-\frac{1}{2}$ these n_{x0} females will have been exposed to the risk of bearing daughters, and the number of the latter they will have given birth to per female alive will be given by the maternal frequency figure for the ages $x+\frac{1}{2}$ to $x+1$. This figure may be obtained by interpolating in the integral curve of the m_x values, and thus expressing the latter in $\frac{1}{2}$ units of age throughout the reproductive span instead of in single units. The daughters born during the interval of time $0-\frac{1}{2}$ will be aged $\frac{1}{2}-1$ at $t = 1$, the number of them surviving at this time being determined approximately by multiplying the appropriate $m_{x+\frac{1}{2}}$ figure by the factor $2\int_{\frac{1}{2}}^{1} l_x dx$ according to the given life table. Similarly, each of the $P_x n_{x0}$ females during the interval of time $\frac{1}{2}-1$ give birth to $m_{x+1-x+1\frac{1}{2}}$ daughters, the survivors of which form part of the $0-\frac{1}{2}$ age group at $t = 1$. The survivorship factor is in this case taken to be $2\int_0^{\frac{1}{2}} l_x dx$.

Combining these two steps together we obtain a series of F_x figures, which may be defined as the number of daughters alive in the age group 0–1 at $t = 1$ per female alive in the age group x to $x+1$ at $t = 0$. Putting

$$k_1 = 2\int_0^{\frac{1}{2}} l_x dx, \quad k_2 = 2\int_{\frac{1}{2}}^{1} l_x dx,$$

then

$$F_x = (k_2 m_{x+\frac{1}{2}-x+1} + k_1 P_x m_{x+1-x+1\frac{1}{2}}),$$

and

$$\sum_{x=0}^{m} F_x n_{x0} = n_{01},$$

the total number of daughters alive aged 0–1 at $t = 1$.

3. Numerical example

In order to see whether the P_x and F_x figures obtained in this way from the basic data give a reasonably accurate estimate of the population in one unit's time, a numerical example was worked out for an imaginary rodent population, the species chosen being the brown rat, *Rattus norvegicus*. Full details of the basic life table and fertility table which were used are given in an appendix, together with a short account of the genesis of these tables and the methods employed to estimate the rate of natural increase (r) and the stable age distribution. Compared with man, the fertility of this imaginary rat population was relatively very great; thus, the gross reproduction rate was 31·21 daughters and the net rate (R_0) 25·66, the life table used being a reasonably good one. The inherent rate of natural increase was estimated to be 0·44565 per head per month of 30 days, and the stable age distribution was so overladen with young that the proportion of females in the post-reproductive age groups was negligible. Some 74·45 % of the females were younger than 3 months, at which age breeding was assumed to commence.

By definition the Malthusian age distribution is stable; that is to say, once a population subject to the given rates of fertility and mortality achieves this form of distribution, it

continues to increase e^r times every unit of time and the proportions of the population alive in each group remain constant. Thus, in the present example, given 100,000 females distributed as to age in the stable form at $t = 0$, the number alive in each age group in 1 month's time can be immediately calculated by multiplying each element in the original distribution by 1·561505. This 'true' age distribution at $t = 1$ is compared in Table 1 with that obtained by operating on the original distribution with the P_x and F_x figures, which are given in Table 5 of the Appendix.

The agreement between the true and estimated age distributions is remarkably close. It might be expected that the principal errors would occur in the early age groups, since the

Table 1

(1) (Units of 30 days) Age group	(2) Population at $t=0$ Stable age distribution	(3) Expected population at $t=1$ Col. 2 × 1·561505	(4) Population at $t=1$ Estimated by operating on col. 2 with the matrix M
0–	37,440	58,463	58,374
1–	22,595	35,282	35,455
2–	14,417	22,512	22,519
3–	9,227	14,408	14,406
4–	5,903	9,218	9,218
5–	3,775	5,895	5,895
6–	2,413	3,768	3,768
7–	1,542	2,408	2,407
8–	984	1,537	1,537
9–	627	979	980
10–	399	623	623
11–	254	397	396
12–	161	251	251
13–	101	158	159
14–	64	100	99
15–	40	62	62
16–	25	39	39
17–	15	23	24
18–	9	14	14
19–	6	9	8
20–	3	5	5
Total	100,000	156,151	156,239

The span of the reproductive ages is from 3 to 21 months.

P_x figures are based on the stationary age distribution which is clearly very different from the stable form. However, as will be seen from Table 1, the biggest error from this cause is due to the first P_0 which overestimates the number alive in the 1–2 age group at $t = 1$ by some 0·5 %. The F_x figures underestimate the number alive in the 0–1 group by 0·2 %, and the total population is overestimated by 0·06 %. On the whole these results are satisfactory and, judging from this example, it would seem that the matrix M operating on a given age distribution should give a reasonable estimate of the population in one unit's time, provided that the unit of time and age chosen be not too coarse as compared with the life span of the species. The degree of cumulative error which is introduced by continued operation with the matrix will be considered later.

4. PROPERTIES OF THE BASIC MATRIX

The matrix M is square and of order $m+1$; it is not necessary, however, in what follows to consider this matrix as a whole. For, if $x = k$ is the last age group within which reproduction occurs, F_k is the last F_x figure which is not equal to zero. Then, if the matrix be partitioned symmetrically at this point,

$$M = \begin{bmatrix} A & . \\ B & C \end{bmatrix}.$$

The submatrix A is square; B is of order $(m-k)\times(k+1)$; C again is square consisting of $m-k$ rows and columns, the only numerical elements being in the subdiagonal immediately below the principal diagonal. The remaining submatrix is of order $(k+1)\times(m-k)$ and consists only of zero elements. Then in forming the series of matrices M^2, M^3, M^4, etc.,

$$M^t = \begin{bmatrix} A^t & . \\ f(ABC) & C^t \end{bmatrix}.$$

The submatrix C is, however, of such a type that $C^{m-k} = 0$, so that M^t, $t \geqq m-k$, will have all its last $m-k$ columns consisting of zero elements. This is merely an expression of the obvious fact that individuals alive in the post-reproductive ages contribute nothing to the population after they themselves are dead. It is the submatrix A which is principally of interest, and in the mathematical discussion which follows, attention is focused almost entirely on it and on age distributions confined to the prereproductive and reproductive age groups.

The matrix A is of order $(k+1)\times(k+1)$, where $x = k$ is the last age group in which reproduction occurs, and written in full,

$$A = \begin{bmatrix} F_0 & F_1 & F_2 & F_3 & \cdots & F_{k-1} & F_k \\ P_0 & . & . & . & \cdots & . & . \\ . & P_1 & . & . & \cdots & . & . \\ . & . & P_2 & . & \cdots & . & . \\ \multicolumn{7}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ . & . & . & . & \cdots & P_{k-1} & . \end{bmatrix}.$$

This matrix is non-singular, since the determinant $|A| = (-1)^{k+2}(P_0P_1P_2\ldots P_{k-1}F_k)$. There exists, therefore, a reciprocal matrix of the form

$$A^{-1} = \begin{bmatrix} . & P_0^{-1} & . & . & \cdots & . \\ . & . & P_1^{-1} & . & \cdots & . \\ . & . & . & P_2^{-1} & \cdots & . \\ \multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ . & . & . & . & \cdots & P_{k-1}^{-1} \\ F_k^{-1} & -(P_0F_k)^{-1}F_0 & -(P_1F_k)^{-1}F_1 & -(P_2F_k)^{-1}F_2 & \cdots & -(P_{k-1}F_k)^{-1}F_{k-1} \end{bmatrix}$$

Thus, given an initial age distribution n_{x0} $(x = 0, 1, 2, 3, \ldots, k)$ at $t = 0$, in addition to the forward series of operations An_0, A^2n_0, A^3n_0, ..., etc., there is also a backward series $A^{-1}n_0$, $A^{-2}n_0$, $A^{-3}n_0$, ..., etc. There is, however, a fundamental difference between these; for, whereas the forward series can be carried on for as long as we like, given any initial age distribution, the backward series can only be performed so long as n_{xt} remains $\geqq 0$, since a negative number of individuals in an age group is meaningless. Apart from this limitation, it is possible to foresee that the reciprocal matrix might be of some use in the solution of certain types of problem.

5. Transformation of the co-ordinate system

Hitherto an age distribution n_{xt} has been regarded as a matrix consisting of a single column of elements. For simplicity in notation, this column vector will now be termed the vector ξ and different ξ's will be distinguished by different subscripts (ξ_a, ξ_x, etc.). We may picture an age distribution as a vector having a certain magnitude and related to a definite direction in a vector space, the space of the ξ's. The different age distributions which may arise in the case of any particular population will be assumed to be ξ's all radiating from a common origin. The numerical elements of a ξ vector are thus taken to be the co-ordinates of a point in multi-dimensional space referred to a general Cartesian co-ordinate system, in which the reference axes may make any angles with one another. At this point in the argument another type of vector will be introduced, which in matrix notation will be written as a row vector, and which will be termed the vector η. There is an intimate relationship between this new type and the old, for, associated with each vector ξ_a, there is a uniquely determined vector η_a, and vice versa. The inner or scalar product, $\eta_a \xi_a$, is the square of the length of the vector ξ_a. Either we may picture each of these vectors as associated with a different kind of vector space, the space of the ξ's and the dual space of the η's, which are not entirely disconnected but related in a special way; or, alternatively, we may regard them as two different kinds of vector associated with the same vector space. The relationship between η and ξ is precisely the same as that between covariant and contravariant vectors in differential geometry.

If we pass from our original co-ordinate system to a new frame of reference, and the variables η and ξ undergo the non-singular linear transformations,

$$\eta = \phi H, \quad \xi = H^{-1}\psi, \quad |H| \neq 0,$$

it can be seen that since the variables are contragredient, $\eta\xi = \phi\psi$, so that the square of the length of a vector remains invariant. Moreover, since the result of operating on a vector ξ_a with the matrix A is, in general, another vector ξ_b, where ξ_a and ξ_b are both referred to the original co-ordinate system, it follows that in the new frame of reference which is defined by the linear transformations given above, the relationship

$$A\xi_a = \xi_b$$

becomes

$$HAH^{-1}\psi_a = \psi_b,$$

or

$$B\psi_a = \psi_b.$$

Thus, in the new frame of reference the matrix $B = HAH^{-1}$ operating on the vector ψ_a is equivalent to the matrix A operating on the vector ξ_a in the original frame.

It is convenient, for the purposes of studying the matrix A and of performing any numerical computations with it, to transform the variables η and ξ in the above way, choosing the matrix H so as to make $B = HAH^{-1}$ as simple as possible. For $B^t = (HAH^{-1})^t = HA^tH^{-1}$ and since A is non-singular, by the reversal law, $(HAH^{-1})^{-1} = HA^{-1}H^{-1}$. Thus, if $f(A)$ is a rational integral function of A, $f(B) = f(HAH^{-1}) = Hf(A)H^{-1}$; and the properties of matrix functions $f(A)$ can be studied by means of the simpler forms $f(B)$. Moreover, the matrices A and B have the same characteristic equation and, therefore, the same latent roots. For $B - \lambda I = H(A - \lambda I)H^{-1}$ and, forming the determinants of both sides,

$$|B - \lambda I| = |H|\,|A - \lambda I|\,|H|^{-1},$$

so that the characteristic equation is

$$|A - \lambda I| = |B - \lambda I| = 0.$$

If, in the present case, the transforming matrix is taken to be

$$H = \begin{bmatrix} (P_0P_1P_2\ldots P_{k-1}) & . & . & \ldots & . & . & . \\ . & (P_1P_2P_3\ldots P_{k-1}) & . & \ldots & . & . & . \\ . & . & (P_2P_3\ldots P_{k-1}) & \ldots & . & . & . \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ . & . & . & \ldots & (P_{k-2}P_{k-1}) & . & . \\ . & . & . & \ldots & . & P_{k-1} & . \\ . & . & . & \ldots & . & . & 1 \end{bmatrix}$$

in which, it is to be noted, the only numerical elements lie in the principal diagonal and are derived entirely from the life table, then

$$B = HAH^{-1} = \begin{bmatrix} F_0 & P_0F_1 & P_0P_1F_2 & P_0P_1P_2F_3 & \ldots & (P_0P_1P_2\ldots P_{k-1})F_k \\ 1 & . & . & . & \ldots & . \\ . & 1 & . & . & \ldots & . \\ . & . & 1 & . & \ldots & . \\ . & . & . & 1 & \ldots & . \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ . & . & . & . & \ldots\ 1 & . \end{bmatrix}.$$

Comparing this matrix B with the original form A, it can be seen that the latter has been simplified to the extent that the original P_x figures in the principal subdiagonal are now replaced by a series of units, and the matrix A has been reduced to the rational canonical form $B = HAH^{-1}$ (see Turnbull & Aitken, 1932, chap. v). In this way any computations with the matrix A are made easier, and we may work henceforward in terms of ϕ and ψ vectors together with the matrix B, instead of with the original η and ξ vectors, and the matrix A. Any results obtained in this new system of co-ordinates may be transformed back again to the original system whenever necessary. It is evident that by suitably enlarging H the original matrix M may be transformed in a similar way.

This linear transformation of the original co-ordinate system is equivalent biologically to the transformation of the original population we were considering into a new and completely imaginary type which, although intimately connected with the old, has certain quite different properties. Thus, it can be seen from the transformed matrix B that the individuals in this new population, instead of dying off according to age as the original ones did, live until the whole span of life is completed, when they all die simultaneously. This is indicated by the P_x figures being now all equal to unity; an individual alive in the age group x to $x+1$ at $t = 0$ is certain of being alive at $t = 1$, excepting in the last age group of all where none of the individuals will be alive in one unit's time. Accompanying this somewhat radical change in the life table, there is a compensatory adjustment made in the rates of fertility so that the new population has the same inherent power of natural increase (r) as that of the old. This follows from the fact that the latent roots of the matrices A and B are the same, and, as will be shown later, the dominant latent root is closely related to the value of r obtained by the usual methods of computation. Insomuch as the transformation is reversible and $A = H^{-1}BH$, it can be seen that by changing H we could transform the canonical form B, if we wished, into another matrix in which the P_x subdiagonal might be a specified set of

figures derived from some other form of life table. But, for our present purposes, the canonical form B, in which all the P_x figures are units, offers advantages over any other matrix of a similar type owing to the greater ease with which it can be handled.

6. Relation between the canonical form B and the $L_x m_x$ column

The actual computation of the matrix B by way of the steps indicated in the theoretical development is by no means difficult, although it is a somewhat tedious process, particularly if the matrix is of a large order. The numerical elements in the first row of B for the brown rat are given in Table 5 of the Appendix. These values were obtained from the F_x and P_x figures which have already been used in the numerical example in § 3 and which will be found in the same table. Further reflection suggested, however, that instead of first of all obtaining A and then transforming to B, a short cut could be taken which would save labour and which also would tend to eliminate some of the small cumulative errors arising in the longer method.

The series of values P_0, P_0P_1, $P_0P_1P_2$, ..., $(P_0P_1P_2 \dots P_{k-1})$ by which the individual F_x figures are multiplied in order to obtain the first row of B, is essentially a stationary age distribution. For, since by definition,

$$P_x = \frac{L_{x+1}}{L_x},$$

$$(P_0P_1P_2 \dots P_x) = \frac{L_{x+1}}{L_0},$$

where $L_0 = \int_0^1 l_x dx$. Hence the required series of multipliers is given by a stationary age distribution in which only one individual is alive in the age group 0–1. Now, the F_x figures, as defined in § 2, already contain within them some allowance not only for the probability of survival during the first unit of life, but also for the fact that some adult individuals in each age group are dying off during the interval of time 0–1. The process of multiplying F_x by $(P_0P_1P_2 \dots P_{x-1})$ is thus analogous to the formation of the $L_x m_x$ column, by means of which the net reproduction rate is estimated. The chief difference between the first row of B and the $L_x m_x$ distribution is that in the former the maternal frequency is expressed as between the ages of $x+\frac{1}{2}$ to $x+1\frac{1}{2}$, instead of between x to $x+1$ as in the latter. If each element $(P_0P_1P_2 \dots P_{x-1}F_x)$ of the first row of B is regarded as centred at the age of $x+1$, the sum, mean and seminvariants of this 'distribution' may be estimated and compared with the values which are obtained from the $L_x m_x$ column in the process of calculating r by the usual methods. In the present numerical example the results of this comparison were as follows:

Parameter	$L_x m_x$ column	First row of B
Sum (R_0)	25·65786	25·6603
Mean	9·60604	9·5948
m_2	14·14397	14·1839
m_3	22·15696	21·9358
$m_4-3m_2^2$	−117·6480	−117·920

After allowing for the small cumulative errors which might be expected to occur in the calculation of the matrix elements, there is a substantial agreement between the respective

estimates. This agreement strongly suggests that if we had wished to pass immediately to the matrix B without going through the laborious process of calculating the F_x and P_x figures, the elements of the first row could have been obtained by forming a new $L_x m_x$ column in which the age group limits were shifted a half unit later in life. This could readily be done by interpolating in the integral curve of the $L_x m_x$ values for the ages $x+\frac{1}{2}$. This method of forming the first row of B has been adopted in other instances, when the matrix A was not of any immediate interest. It proved to be relatively quick and certainly less laborious than the method of first establishing A and then transforming to B which was the one used in the present numerical example.

7. THE STABLE AGE DISTRIBUTION

The result of operating on an age distributon ψ_x with the matrix B is, in general, a different distribution ψ_y. But, in the special case when the relation between the two distributions is such that

$$B\psi_a = \lambda\psi_a,$$

where λ is an algebraic number, then ψ_a may be said to be a stable age distribution appropriate to the matrix B. For the sake of brevity it will be referred to as a stable ψ. Similarly for initial row vectors, if

$$\phi_a B = \lambda\phi_a,$$

then ϕ_a is said to be a stable ϕ.

The matrix equation defining a stable ψ may be written as $k+1$ linear equations, of which the ith is

$$\sum_{j=1}^{k+1} b_{ij} n_j - \lambda n_i = 0,$$

where n_i $(i = 1, 2, \ldots, k+1)$ are the co-ordinates of the stable ψ, and b_{ij} the element in the ith row and jth column of B. Eliminating the n_i from this system of equations, we obtain the characteristic equation of B, namely,

$$|B-\lambda I| = 0;$$

and, expanding this determinant in powers of λ, we have in the present case,

$$\lambda^{k+1} - F_0\lambda^k - P_0F_1\lambda^{k-1} - P_0P_1F_2\lambda^{k-2} - \ldots - (P_0P_1\ldots P_{k-2})F_{k-1}\lambda - (P_0P_1\ldots P_{k-1})F_k = 0.$$

The $k+1$ roots λ_a of this equation are the latent roots of B, and corresponding to each distinct λ_a there is a pair of stable vectors, ϕ_a and ψ_a, determined except for an arbitrary scalar factor.

Once a latent root λ_a has been determined, it is a comparatively simple matter to find the appropriate stable ψ_a and ϕ_a vectors. Thus, it is easily shown that the stable ψ_a is the column vector $\{\lambda_a^k \lambda_a^{k-1} \lambda_a^{k-2} \ldots \lambda_a 1\}$. A short method of estimating ϕ_a is the following. Suppose, to take a simple case, that

$$B = \begin{bmatrix} a & b & c & d \\ 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \end{bmatrix}$$

and let y_x $(x = 1, 2, 3, 4)$ be the elements of the stable ϕ_a appropriate to the root λ_a. Then

$$\begin{aligned}\phi_a B &= [ay_1+y_2 \quad by_1+y_3 \quad cy_1+y_4 \quad dy_1] \\ &= [\lambda_a y_1 \quad \lambda_a y_2 \quad \lambda_a y_3 \quad \lambda_a y_4].\end{aligned}$$

By equating similar elements and putting $y_1 = 1$, $y_4 = d/\lambda_a$, $y_3 = \dfrac{c+y_4}{\lambda_a}$, etc., it is easy to see how the required row vector can be built up. Having in this way obtained the stable ψ and ϕ vectors for the matrix B, they may be transformed to the appropriate stable ξ and η for the matrix A by means of the relations

$$\eta = \phi H, \quad \xi = H^{-1}\psi.$$

The characteristic equation of the matrix B, when expanded, is of degree $k+1$ in λ, and once B has been obtained this equation can immediately be written down, since the numerical coefficients of λ^k, λ^{k-1}, λ^{k-2}, etc., are merely the elements of the first row taken with a negative sign. Since there is only one change of sign in this equation, only one of the latent roots will be real and positive. Excluding the rather special case when the first row of B has only a single non-zero element, and taking the more usual type of matrix which will be met with, namely, that for a species breeding continuously over a large proportion of its total life span, it will be found that the modulus of this root (λ_1) is greater than any of the others,

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots > |\lambda_{k+1}|,$$

the remaining roots being either negative or complex.

This dominant latent root λ_1, which will be $\gtreqless 1$ according as to whether the sum of the elements in the first row of B is $\gtreqless 1$, is the one which is principally of interest. Since it is real and positive, it is the only root which will give rise to a stable ψ or ξ vector consisting of real and positive elements. It is this stable ξ_1 associated with the dominant root λ_1 which is ordinarily referred to as the stable age distribution appropriate to the given age specific rates of fertility and mortality. Since

$$A^t\xi_1 = \lambda_1^t\xi_1,$$

it can be seen that the latent root λ_1 of the matrix A and the value of r obtained in the usual way from

$$\int_0^\infty e^{-rx} l_x m_x \, dx = 1,$$

are related by

$$\log_e \lambda_1 = r.$$

From the mathematical point of view, however, the negative and complex roots of the characteristic equation are of importance in the further theoretical development. Moreover, as will be shown later, the stable vectors associated with them are not entirely without interest. Two main cases then arise: when the remaining roots are all distinct, and when there are repeated roots. For the present it will be assumed that the latent roots of the matrix are all distinct.

8. Properties of the stable vectors

Before proceeding further it is necessary to mention briefly the reasons why the methods given above for the computation of the stable ψ and ϕ vectors were adopted, apart from their simplicity in practice. If the $k+1$ distinct roots of the characteristic equation are known, we may form a set of $k+1$ matrices $f(\lambda_a)$ by inserting in turn the numerical value of each root in the matrix $[B - \lambda_a I]$. The adjoint of $f(\lambda_a)$ is

$$F(\lambda_a) = \prod_{b \neq a} [B - \lambda_b I] \quad \text{and} \quad f(\lambda_a) F(\lambda_a) = 0.$$

It may be shown that the stable ψ_a appropriate to the root λ_a can be taken proportional to any column, and the stable ϕ_a proportional to any row of the matrix $F(\lambda_a)$ (see e.g. Frazer, Duncan & Collar, 1938, chap. III). Moreover, $F(\lambda_a)$ is a matrix product of the type $\psi\phi$,

where the ψ vector is given by the first column and the ϕ vector by the last row of $F(\lambda_a)$, each divided by the square root of the element in the bottom left-hand corner; and the trace of the matrix is equal to the scalar product $\phi\psi$. Now $[B-\lambda_a I]$ is a square matrix of order $k+1$ with only zero elements below and to the left of the principal subdiagonal, which itself consists of units. The product of k such matrices, which gives $F(\lambda_a)$, will have therefore a unit in the bottom left-hand corner. Since the stable ϕ_a and ψ_a vectors obtained by the methods suggested in § 7 have respectively their first and last elements $= 1$, it follows that

$$\psi_a\phi_a = F(\lambda_a), \quad \phi_a\psi_a = \text{trace}\, F(\lambda_a).$$

The stable vectors may now be normalized. If the scalar product, $\phi_a\psi_a = z^2$, say, then

$$\frac{\phi_a}{|z|}\frac{\psi_a}{|z|} = 1.$$

From now on it will be assumed that the stable vectors appropriate to each of the latent roots have been normalized in this way.

These vectors have the following important properties:

(1) The $k+1$ stable ψ are linearly independent. There is thus no such relationship, with non-zero coefficients c, as

$$c_1\psi_1+c_2\psi_2+c_3\psi_3+\ldots+c_{k+1}\psi_{k+1} = 0.$$

(2) The scalar product of a stable ψ, ψ_a with the associated vector of another stable ψ, ψ_b is zero, i.e.

$$\phi_b\psi_a = 0 \quad (a \neq b).$$

The normalized stable ψ thus form a set of $k+1$ independent and mutually orthogonal vectors of unit length.

(3) Any arbitrary ψ—ψ_x say—can be expanded in terms of the stable ψ, thus

$$\psi_x = c_1\psi_1+c_2\psi_2+c_3\psi_3+\ldots+c_{k+1}\psi_{k+1},$$

where the coefficients c may be either real or complex. Similarly an arbitrary vector ϕ_x can be expanded in terms of the stable ϕ.

9. The spectral set of operators

The matrix product $\psi_a\phi_a$ of the normalized stable vectors associated with the latent root λ_a will be termed the matrix S_a. From the relationships which have already been given, it can be seen that S_a is merely the adjoint matrix $F(\lambda_a)$ of the previous section after each element in the latter has been divided by the sum of the elements in the principal diagonal; in other words it is the normalized $F(\lambda_a)$. In the case of all the latent roots being distinct, there are thus $k+1$ matrices S_a, and these S_a form a spectral set of operators with the following properties:

$$S_a^2 = S_a, \quad S_aS_b = 0 \quad (a \neq b), \quad \sum_{a=1}^{k+1} S_a = I.$$

Moreover, if $f(B)$ is a polynomial of the matrix B, we have by Sylvester's theorem (Turnbull & Aitken, 1932, chap. VI, §8)

$$f(B) = \sum_{a=1}^{k+1} f(\lambda_a)\, S_a,$$

so that the matrix

$$B = \lambda_1 S_1+\lambda_2 S_2+\ldots+\lambda_{k+1}S_{k+1}, \quad \text{and} \quad B^t = \lambda_1^t S_1+\lambda_2^t S_2+\ldots+\lambda_{k+1}S_{k+1}.$$

If the latent roots in the expansion of B are raised to a high power, the term associated with the positive real root predominates over all the others, so that when t is large, we have approximately

$$B^t = \lambda_1^t S_1.$$

In any particular case the power to which B will have to be raised in order that this equation should be approximately true, will depend both on the order of the matrix and on the relative magnitude of the dominant root as compared with that of the remaining roots.

At this point it is possible to attach some biological meaning to one of the ϕ or η row vectors, which in the first place were introduced into the theory for reasons of symmetry, and which were defined solely in terms of their mathematical properties. If at a given moment a transformed population has an arbitrary age distribution ψ_x, and the sequence $B\psi_x, B^2\psi_x, \ldots, B^t\psi_x$ is formed, it can be seen that when t is large and ψ_x is expanded in terms of the stable ψ, we have approximately

$$B^t\psi_x = c_1\lambda_1^t\psi_1.$$

Thus, a population with any arbitrary age distribution tends ultimately to approach the stable form appropriate to the given rates of fertility and mortality, provided that these age-specific rates remain constant. This theorem is, of course, well known; and it is clear that the achievement of the stable form of age distribution associated with the dominant latent root is very unlikely to occur in practice, except in the case when the initial distribution is already of that form or exhibits only small departures from it. Now, it has already been shown that the sums of the columns of a matrix B^t provide a measure of the contributions made to the population at time t per individual alive in the respective age groups at $t = 0$. When t is large, the matrix B^t is equivalent to the matrix S_1 multiplied by a scalar factor. From the way in which this latter matrix was constructed by the outer multiplication of ψ_1 and ϕ_1, it is evident that the sums of the columns of S_1 are proportional to the vector ϕ_1. Thus, transforming back again to the original co-ordinate system, the stable η_1 associated with the dominant latent root provides a measure of the relative contributions per head made to the stable population by the individual age groups.

10. Reduction of B to classical canonical form

From the $k+1$ stable ψ a matrix Q can be constructed, whose columns are the stable ψ arranged, reading from left to right, in descending order of the moduli of the roots with which they are associated. Corresponding to every pair of complex roots, $u \pm iv$, there will be in this matrix a pair of columns consisting of complex elements, the one column being the conjugate complex of the other. Some of the columns associated with the negative roots may be purely imaginary owing to the normalization of the corresponding ψ and ϕ vectors. In a similar way a matrix U may be formed, whose rows, reading from above down, are the stable ϕ arranged in the same order. Since the stable ϕ and ψ are normalized, and $\phi_a\psi_b = 0$ for $a \neq b$,

$$UQ = I,$$

and, therefore, U and Q are reciprocal matrices. By premultiplying and postmultiplying respectively with U and Q, the matrix B may be reduced to the classical canonical form C, in which the only elements lie in the principal diagonal and consist of the latent roots arranged in the order prescribed above. This reduction of B to a purely diagonal form by means of the collineatory transformation $UBQ = C$ is, however, only possible in the type of matrix we are considering, when the latent roots are all distinct.

The expansion of an arbitrary ψ_x in terms of the stable ψ,

$$\psi_x = c_1\psi_1 + c_2\psi_2 + \ldots + c_{k+1}\psi_{k+1},$$

may be written in matrix notation as $\psi = Qc$,

where c is the column vector $\{c_1c_2c_3\ldots c_{k+1}\}$. Similarly, the expansion of the vector ϕ_x associated with ψ_x may be written $\phi = dU$,

where d is the row vector $[d_1d_2d_3\ldots d_{k+1}]$. This is again a transformation to another co-ordinate system, but this time the reference axes are at right angles to one another. Since the variables transform contragrediently, $dc = \phi\psi = \eta\xi$. At this point it is necessary to make some assumption as to the relationship between the elements of the vectors d and c. Since these elements may be either real or complex, it will be assumed that

$$d = \bar{c}',$$

where the row vector $\bar{c}'$ is the transposed conjugate complex of the column vector c. Hence, *the square of the length of a vector referred to this orthogonal co-ordinate system is given by* $\bar{c}'c$, a number which is essentially real and non-negative. (The assumption that $d = c'$ will be found, in the particular case studied here, to lead to values of $c'c$ which, although real, may be negative.)

11. The relation between ϕ and ψ vectors

Since $\psi_x = Qc_x$, and the associated $\phi_x = \bar{c}'_x U$, it may be seen from the relations given in the two previous sections that $\bar{\phi}'_x = \bar{U}'U\psi_x = G\psi_x$.

The matrix $G = \bar{U}'U$ is symmetrical and all its elements are real numbers, those in the principal diagonal being necessarily positive in sign. It therefore remains unaltered after transposition. Since the elements of the vector ψ_x, which is by definition an age distribution transformed by the matrix H, are also necessarily real, we may write

$$\phi_x = \psi'_x G.$$

The role of the matrix G is therefore the same as that of the double covariant metric tensor g_{mn} in the tensor calculus. It transforms any ψ vector into its associated ϕ vector. This process is reversible, the reciprocal matrix being given by $G^{-1} = Q\bar{Q}'$.

The magnitude of a vector ψ_x is defined by the equation

$$x = (\psi'_x G\psi_x)^{\frac{1}{2}},$$

where the square root is taken with a positive sign. If we have two vectors ψ_x and ψ_y, both radiating from the common origin, the angle between them is given by

$$\cos\theta = \frac{\psi'_x G\psi_y}{xy},$$

from which it follows that when $\psi'_x G\psi_y = \phi_x\psi_y = 0$, the two vectors are at right angles to one another, and when $\cos\theta = 1$ their directions are the same. If, in the last equation, we take ψ_y to be the stable vector ψ_1 associated with the dominant latent root λ_1, then knowing the magnitude of a vector $B^t\psi_x$ and the angle which it makes with the ψ_1 axis, we can obtain a graphical representation of the way in which a particular age distribution approaches the stable form

The matrix G also defines the angles between the reference axes of the co-ordinate system. If we introduce into the vector space of the ψ's a system of reference axes defined by the unit column vectors

$$\begin{aligned} e_1 &= \{1 \quad 0 \quad 0 \quad \dots \quad 0\} \\ e_2 &= \{0 \quad 1 \quad 0 \quad \dots \quad 0\} \\ e_3 &= \{0 \quad 0 \quad 1 \quad \dots \quad 0\} \\ &\vdots \\ e_{k+1} &= \{0 \quad 0 \quad 0 \quad \dots \quad 1\} \end{aligned}$$

then the distance from the origin of the unit point e_j is

$$Oe_j = \sqrt{g_{jj}},$$

and the angle between any two of the co-ordinate axes is given by

$$\cos\theta_{ij} = \frac{g_{ij}}{\sqrt{(g_{ii}g_{jj})}},$$

where, in both cases, g_{ij} is the element in the ith row and jth column of G.

By transforming back to the original co-ordinate system, the metric matrix associated with the vector space of the ξ's will be found to be

$$G_\xi = HG_\psi H.$$

Hitherto we have been chiefly concerned with an operator B which, acting in the vector space of the ψ's, has the power of transforming a vector ψ_x into what is in general a new vector ψ_y. We may now have reason to inquire how the associated vectors ϕ_x and ϕ_y are related in the vector space of the ϕ's, whenever

$$B\psi_x = \psi_y.$$

Since

$$\psi_x = G^{-1}\phi'_x \quad \text{and} \quad \psi_y = G^{-1}\phi'_y,$$

we have

$$GBG^{-1}\phi'_x = \phi'_y,$$

and hence, by transposing,

$$\phi_x G^{-1}B'G = \phi_y.$$

Thus, the matrix which transforms ϕ_x into ϕ_y is not the same as that which transforms ψ_x into ψ_y. In order to distinguish these two operators, they will be referred to as B_ϕ and B_ψ respectively. In the few numerical examples which have been worked out, the matrix B_ϕ differed greatly from the rational canonical form B_ψ, and consisted of $(k+1)^2$ real elements, some of which were negative. In addition to the relationship

$$B_\phi = G^{-1}B'_\psi G,$$

it was also found that in the case of distinct latent roots

$$B_\phi = \bar{\lambda}_1 S_1 + \bar{\lambda}_2 S_2 + \bar{\lambda}_3 S_3 + \dots + \bar{\lambda}_{k+1} S_{k+1},$$

where the S_a matrices are the spectral set of operators defined in § 9 and $\bar{\lambda}_a$ is the conjugate complex of the latent root λ_a. It may be seen from this expansion of B_ϕ that the necessary condition for $B_\phi = B_\psi$ is that all the latent roots of B should be real, the one positive and the remainder negative. (It is to be noted that we are dealing here with the case of distinct latent roots; it would appear that even if all the λ were real, $B_\phi \neq B_\psi$ in the case of repeated roots.) Unless, however, the matrix B is of a small order, it is unlikely that this condition would be fulfilled, since in the more usual-sized matrix we shall be dealing with in the case

of human or other mammalian populations, some of the roots will almost certainly be complex.*

It seems unlikely that the equations given in this section will have very much practical application at the moment; they have been included merely to fill in the picture of the relationship between the two types of vector. For all ordinary purposes no one would choose to work in terms of ϕ vectors and the operator B_ϕ, instead of the more obvious ψ vectors and the more simple matrix form B_ψ. Nevertheless, since it has been necessary to assume that there are such vectors as η or ϕ associated with every ξ or ψ, and since these vectors play such an important part in the mathematical theory, the question naturally arises as to what significance must be attached to them from the biological point of view. Have they in fact any real meaning at all? Or must they be regarded purely as mathematical abstractions? At the end of § 9 it has been suggested that the row vector associated with the stable ξ_1 appropriate to the dominant latent root is a measure of the contributions made to the stable population per individual female alive in the respective age groups of the initial distribution; but this is a special case and the interpretation offered here is not applicable, even in a wider form, to η vectors in general. It may well be, of course, that the latter as a class have no concrete meaning: and that in seeking to define them in terms of some property or characteristic of an age distribution one is merely attempting the impossible. But the fact of one η vector having been defined in non-mathematical terms, even though on further consideration some revision may be needed of the actual definition given here, suggests that impossible may perhaps be too final a word to use in this connexion.

12. Case of repeated latent roots

When any of the latent roots other than the real positive dominant root are repeated, a number of the relations given in the previous sections no longer hold good and certain equations must therefore be modified. Suppose a root λ_a has a multiplicity s, and consider the matrix $f(\lambda_a)$ such as, to take a simple example,

$$f(\lambda_a) = \begin{bmatrix} a-\lambda_a & b & c & d \\ 1 & -\lambda_a & 0 & 0 \\ 0 & 1 & -\lambda_a & 0 \\ 0 & 0 & 1 & -\lambda_a \end{bmatrix}.$$

Then, since the determinant $|f(\lambda_a)| = 0$ and at least one of the first minors of order 3 is not equal to zero, the above matrix has rank 3 and, therefore, nullity 1. Hence it can be seen that $f(\lambda_a)$ of whatever order it may be has nullity 1. Since $f(\lambda_a)$ is thus simply degenerate, there is only one stable ψ appropriate to the s equal roots λ_a (see e.g. Frazer, Duncan & Collar, 1938, chap. III).

Certain consequences immediately follow. Since the matrices Q and U cannot be constructed in the way given in § 10, the reduction of B to a purely diagonal matrix by means of the collineatory transformation UBQ can no longer be carried out. Neither is the expansion of B in terms of the spectral set of S_a matrices, nor the expansion of an arbitrary ψ_x

* The interesting theoretical case of the matrix A or B having a number of its latent roots real and positive, with the remainder real and negative, is outside the scope of the present study. The necessary conditions for this to be true would involve a number of the F_x figures becoming negative, a case not considered here, but which biologically might be held to correspond with the destruction of eggs, or the very young, by certain age groups, e.g. as observed by Chapman (1933) in experimental populations of the flour beetle, *Tribolium confusum*.

in terms of the stable ψ, possible in the actual forms given in §§ 9 and 8. We may, however, proceed as follows.

The matrix Q is essentially an alternant which has been postmultiplied by a diagonal matrix N, the elements in the latter being given by the reciprocals of the scalar factors $|z|$ by which the stable vectors were divided in the process of normalization $\left(\frac{\phi_a}{|z|}\frac{\psi_a}{|z|}=1\right)$, so that $Q = XN$, where

$$X = \begin{bmatrix} \lambda_1^k & \lambda_2^k & \dots & \lambda_{k+1}^k \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \dots & \lambda_{k+1}^{k-1} \\ \vdots & \vdots & & \vdots \\ \lambda_1 & \lambda_2 & \dots & \lambda_{k+1} \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

When a root λ_a is repeated s times, s of the columns in X become the same, and therefore the matrix becomes singular. In place of this alternant matrix we have the confluent alternant form (see Turnbull & Aitken, 1932, chap. VI) in which the s columns ($s = 0, 1, 2, 3, \dots, s-1$) corresponding to the repeated root λ_a are written

$$\begin{matrix} \lambda_a^k & k\lambda_a^{k-1} & \frac{k(k-1)\,\lambda_a^{k-2}}{2\,!} & \dots \\ \lambda_a^{k-1} & (k-1)\,\lambda_a^{k-2} & \frac{(k-1)\,(k-2)}{2\,!}\lambda_a^{k-3} & \dots \\ \vdots & \vdots & \vdots & \dots \\ \vdots & \vdots & 3\lambda_a & \dots \\ \vdots & 2\lambda_a & 1 & \dots \\ \lambda_a & 1 & 0 & \dots \\ 1 & 0 & 0 & \dots \end{matrix}$$

the column s being obtained from column 0 (the non-normalized stable ψ_a) by the operation $\left(\frac{d}{d\lambda_a}\right)^s \Big/ s!$. This confluent alternant form of X is non-singular and therefore a reciprocal matrix can be determined (see Aitken, 1939, § 50). The general classical canonical form of B obtained by the collineatory transformation, $X^{-1}BX = C$, has corresponding to the repeated root λ_a a diagonal submatrix:

$$\begin{bmatrix} \lambda_a & 1 & . & . & . & . \\ . & \lambda_a & 1 & . & . & . \\ . & . & \lambda_a & 1 & . & . \\ \dots & \dots & \dots & \dots & \dots & \dots \\ . & . & . & . & \lambda_a & 1 \\ . & . & . & . & . & \lambda_a \end{bmatrix}.$$

The matrix product of a column of X with the appropriate row of the reciprocal X^{-1} gives as before S_a and the $k+1$ S_a form a spectral set with the same properties as those defined in § 9, except that

$$B \neq \sum_{a=1}^{k+1} \lambda_a S_a.$$

In place of this expansion of B in terms of the S_a matrices, we have in the case of repeated roots the confluent form of Sylvester's theorem, for details of which reference may be made to Frazer, Duncan & Collar, 1938, chap. III. Apart from this modification, however, we may obtain by inspection of the S_a the factors by which the respective columns of X must be divided in order to express this matrix in a form comparable to that of Q in § 10. Similarly, when the respective rows of X^{-1} are multiplied by the appropriate factors, the matrix U is found and hence $G = \overline{U}'U$ can be constructed. (It is to be noted that $(\overline{X^{-1}})' X^{-1}$ is neither equal to, nor directly proportional to $\overline{U}'U$.) An arbitrary ψ_x can be expanded in terms of the column vectors of Q, though in the case of only one column associated with the repeated root λ_a does the relationship $B\psi_a = \lambda_a \psi_a$ hold.

13. The approach to the stable age distribution

A stable age distribution appropriate to the matrix B has been defined mathematically by the equation

$$B\psi = \lambda\psi,$$

and it has already been shown that since only one latent root of B is real and positive, only one of the stable ψ will consist of real and positive elements. But, in addition to this Malthusian age distribution, it is also of some interest to inquire whether any significance can be attached to the remaining stable ψ associated with the negative and complex roots of the characteristic equation.

Any age distribution ψ_x, the elements of which are necessarily $\geqq 0$, may be expressed as a vector of deviates from the stable ψ_1 associated with the dominant latent root, and we may therefore write the expansion of ψ_x in terms of the stable ψ as

$$(\psi_x - c_1\psi_1) = c_2\psi_2 + c_3\psi_3 + \ldots + c_{k+1}\psi_{k+1} = \psi_d,$$

where the coefficients c are given by the vector $c = U\psi_x$. Thus, the way in which a particular type of age distribution will approach the stable form may be studied by means of the vector ψ_d.

Among the terms occurring in the right-hand side of this expression there will be, corresponding to each negative root, a single term $c_a\psi_a$ which will consist of real elements alternately positive and negative in sign. (Even if the normalized ψ_a is imaginary this term will consist of real numbers, since in this case c_a will also become imaginary.) Moreover, corresponding to every pair of complex roots there will be a pair of terms $(c_m\psi_m + c_n\psi_n)$ which taken together will also give a single vector with real elements. This follows from the fact that c_m is the conjugate complex of c_n owing to the way in which the matrix U is constructed. Then, apart from the scalar c_1 which must necessarily be > 0, some of the remaining coefficients $c_2, c_3, \ldots, c_{k+1}$ in the expansion of ψ_d may be zero. The first and most obvious case is when they are all zero, and the age distribution ψ_x is therefore already of the stable form. But, if either

$$\psi_d = c_a\psi_a,$$

where ψ_a corresponds to a negative latent root, or

$$\psi_d = c_m\psi_m + c_n\psi_n,$$

where ψ_m and ψ_n are associated with a conjugate pair of complex roots, then it follows that the age distribution ψ_x will, as time goes on, approach the stable form in a particular way defined by either

$$B^t\psi_d = c_a\lambda^t\psi_a \quad \text{or} \quad B^t\psi_d = c_m\lambda^t\psi_m + c_n\overline{\lambda}^t\psi_n,$$

in which λ^t for a pair of complex roots $u+iv$ with modulus r may be written in the form of $r^t(\cos\theta t \pm i\sin\theta t)$. Thus, the negative and complex latent roots of B serve to determine a number of age distributions which are of some interest owing to the fact that they will approach the Malthusian form in what may be termed a stable fashion.

Since $|\lambda_1|>|\lambda_2|>|\lambda_3|>\ldots>|\lambda_{k+1}|$, the vector of deviates ψ_d will tend towards zero as $t\to\infty$ whenever $\lambda_1 \lesseqgtr 1$. Thus, in the case of a stationary population, any ψ_x will converge to the stable form of age distribution. But if $\lambda_1>1$, there is a possibility of one or more of the remaining roots having a modulus $\geqq 1$, e.g. $|\lambda_2|\geqq 1$. In the latter case there may be certain age distributions with $c_2 \neq 0$ for which the amplitude of the deviations from the stable form tend either to increase ($|\lambda_2|>1$), or to remain constant ($|\lambda_2|=1$). From the practical point of view, however, we may still say that a population with such an age distribution approaches or becomes approximately equal to the stable population, since λ_1^t is much greater than λ_2^t when t is large.

14. Special case of the matrix with only a single non-zero F_x element

The interesting case of the matrix A having only a single non-zero element in the first row has been illustrated in a numerical example by Bernardelli (1941).* This author has also used a matrix notation in the mathematical appendix to his paper, and the form of his basic matrix is the same as that referred to here as M or A. It is not clear, however, from the definitions which he gives whether he regards the elements in the first row of his matrix as being constituted by the maternal frequency figures (m_x) themselves, or by a series of values similar to those defined here as the F_x figures. He refers to them merely as the specific fertility rates for female births.

In discussing the causes of population waves, Bernardelli describes a hypothetical species, such as a beetle, which lives for only three years and which propagates in the third year of life. He assumes, for the sake of argument, that—to employ the terminology used here—$P_0=\frac{1}{2}$ and $P_1=\frac{1}{3}$, and that 'each female in the age 2–3 produces, on the average, 6 new living females'. Assuming for the moment that he is here defining a F_x figure, we may write this system of mortality and fertility rates as

$$A=\begin{bmatrix}0 & 0 & 6\\ \frac{1}{2} & 0 & 0\\ 0 & \frac{1}{3} & 0\end{bmatrix},\quad B=HAH^{-1}=\begin{bmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}.$$

The characteristic equation expanded in terms of λ is $\lambda^3-1=0$; and the latent roots are therefore $1,\ -\frac{1}{2}\pm\frac{\sqrt{3}}{2}i$, all three being of equal modulus. The matrix A has the interesting properties

$$A^2=A^{-1},\quad A^3=I,$$

so that any initial age distribution repeats itself regularly every three years. Thus, as Bernardelli shows, a population of 3000 females distributed equally among the three age

* At this point I should like to acknowledge the gift of a reprint of this paper, which was received by the Bureau of Animal Population at a time when I was in the middle of this work, and when I was just beginning to appreciate the interesting results which could be obtained from the use of matrices and vectors: also a personal communication from Dr Bernardelli, received early in 1942, at a time when it was difficult to reply owing to the developments of the war situation in Burma. Although the problems we were immediately interested in differed somewhat, this paper did much to stimulate the theoretical development given here, and it is with great pleasure that I acknowledge the debt which I owe to him.

groups becomes a total population of 6833 at $t = 1$; of 5166 at $t = 2$; and again 3000 distributed equally among the age groups at $t = 3$. Unless a population has already an initial age distribution in the ratio of $\{6:3:1\}$, no approach will be made to the stable form associated with the real latent root, and the vector of deviates ξ_d will continue to oscillate with a stable amplitude, which will in part depend on the form of the initial distribution. Although this numerical example refers specifically to a stationary population, it is evident that a similar type of argument may be developed in the case when $|\lambda| > 1$ and $A^3 = \lambda^3 I$.

We have assumed here that his definition of the fertility rate refers to a F_x figure. But, if we were to interpret the words quoted above as referring to a maternal frequency figure, namely that every female alive between the ages 2–3 produces on the average 6 daughters per annum, then the results become entirely different. For, deriving the appropriate F_x figures by the method described in § 2, the matrix is now

$$A = \begin{bmatrix} 0 & 1 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix}, \quad B = HAH^{-1} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

and the latent roots are $1, -\frac{1}{2} \pm \frac{1}{2}i$. The modulus of the pair of complex roots is $1/\sqrt{2}$, which is < 1, so that every age distribution will now converge to the stable form associated with the real root. Thus, to take the same example as before, 3000 females distributed equally among the age groups will tend towards a total population of 4000 distributed in the ratio of $\{6:3:1\}$, and it was found that this age distribution would be achieved at approximately $t = 23$. During the approach to this stable form periodic waves are apparent both in the age distribution and in the total number of individuals, but these oscillations are now damped, in contrast with the results obtained with the first type of matrix.

This simple illustration serves to emphasize the importance which must be attached to the way in which the basic data are defined and to the marked difference which exists between what are termed here the m_x and F_x figures. Nevertheless, apart from the question of the precise way in which the definition of the fertility rates is to be interpreted in this example, the first type of matrix with only a single element in the first row does correspond to the reproductive biology of certain species. Thus, in the case of many insect types the individuals pass the major portion of their life span in various immature phases and end their lives in a short and highly concentrated spell of breeding. The properties of this matrix suggest that any stability of age structure will be exceptional in a population of this type, and that even if the matrix remains constant we should expect quite violent oscillations to occur in the total number of individuals.

15. NUMERICAL COMPARISON WITH THE USUAL METHODS OF COMPUTATION

From the practical point of view it will not always be necessary to estimate the actual values of all the stable vectors and of the associated matrices which are based on them. Naturally, much will depend on the type of information which is required in any particular case. In order to compute, for instance, the matrices U, Q and G, it is necessary first of all to determine all the latent roots of the basic matrix. The ease with which these may be found depends very greatly upon the order of the matrix. Thus, in the numerical example for the brown rat used previously in § 3, the unit of age and time is one month and the resulting square matrix A is of order 21. To determine all the 21 roots of the characteristic equation would be

a formidable undertaking. It might be sufficient in this case to estimate the positive real root and the stable vector associated with it. On the other hand, it is possible to reduce the size of the matrix by taking a larger unit of age, and in some types of problem, where extreme accuracy is not essential, a unit say three times as great might be equally satisfactory, which would reduce the matrix for the rat population to the order of 7×7. It is not too difficult to find all the roots of a seventh degree equation by means of the root-squaring method (Whittaker & Robinson, 1932, p. 106). But the reduction of the matrix in this way will generally lead to a value of the positive real root which is not the same as that obtained from the larger matrix, and it is therefore necessary to see by how much these values may differ owing to the adoption of a larger unit of time.

Another important point which must be considered is the following. By expressing the age specific fertility and mortality rates in the form of a matrix and regarding an age distribution as a vector, an element of discontinuity is introduced into what is ordinarily taken to be a continuous system. Instead of the differential and integral calculus, matrix algebra

Table 2

Age group (units of 30 days)	'True' stable age distribution	Matrix stable age distribution	Age group (units of 30 days)	'True' stable age distribution	Matrix stable age distribution
0–	37,440	37,362	12–	161	160
1–	22,595	22,644	13–	101	101
2–	14,417	14,444	14–	64	63
3–	9,227	9,238	15–	40	40
4–	5,903	5,906	16–	25	24
5–	3,775	3,775	17–	15	15
6–	2,413	2,412	18–	9	9
7–	1,542	1,540	19–	6	5
8–	984	982	20–	3	3
9–	627	626			
10–	399	398			
11–	254	253	Total	100,000	100,000

is used, a step which leads to a great economy in the use of symbols and consequently to equations which are more easily handled. Moreover, many quite complicated arithmetical problems can be solved with great ease by manipulating the matrix which represents the given system of age specific rates. But the question then arises whether these advantages may not be offset by a greater degree of inaccuracy in the results as compared with those obtained from the previous methods of computation. It is not easy, however, to settle this point satisfactorily. In the way the usual equations of population mathematics are solved, a similar element of discontinuity is introduced by the use of age grouping. Thus, in the case of a human population, if we were estimating the inherent rate of increase in the ordinary way, we should not expect to obtain the same value of r from the data grouped in five year intervals of age as that from the data grouped in one year intervals. The estimates of the seminvariants would not be precisely the same in both cases. Nevertheless, the estimate from the data grouped in five year intervals is usually considered to be sufficiently accurate for all ordinary purposes, and there is little doubt that if we merely require the inherent rate of increase and the stable age distribution, these methods of computation are perfectly satisfactory when applied to human data. But, in the case of rodents, and probably also

other species with high gross and net reproduction rates, it will be found that even a 4th degree equation in r with the coefficients based on the seminvariants of the $L_x m_x$ distribution is, in many examples, not sufficient to give an accurate estimate of the rate of increase, and it is necessary to arrive at a better value of r in a somewhat roundabout way. Here, the determination of the positive real root of the characteristic equation for the matrix, once the latter has been established, may be even quicker than finding a solution from the $L_x m_x$ column by a method such as that described in the appendix.

In order to compare the values of r obtained from the characteristic equation of the matrix with those obtained from the $L_x m_x$ column, both methods were used in the numerical example for the brown rat, and a comparison was also made between the values when the data were grouped in 1 month and in 3 month age intervals. In addition, the stable age distribution appropriate to the positive real root of the matrix was also calculated in both cases. The results were as follows.

(a) *One month unit of grouping; matrix of order* 21×21. Using the method of computation indicated in the appendix, the value of r was estimated to be 0·44565 per month of 30 days. The positive real root of the characteristic equation was $\lambda_1 = 1{\cdot}56246$, whence $r = 0{\cdot}44626$, a value which differs from the former only in the fourth decimal place. The appropriate age distributions, expressed per 100,000, are given in Table 2, the 'true' stable being obtained from

$$n_x = 100{,}000b\int_x^{x+1} e^{-rx} l_x \, dx.$$

The agreement between these distributions is very good, although the one derived from the matrix shows certain small rhythmical departures from the 'true' distribution particularly in the earlier age groups. The maximum difference between them in this region, however, is not greater than 2·2 per thousand. Since the matrix stable distribution is proportional to the columns of the matrix $H^{-1}S_1H$, which in turn is proportional to A^t when t is large, this agreement between the two distributions also indicates that the cumulative errors which might be expected in forming the series A^2, A^3, A^4, etc., owing to the P_x figures being based on the life table age distribution, are not very serious. Judging by this example it seems that satisfactory estimates of the inherent rate of increase and of the stable age distribution may be made from a large order matrix.

(b) *Three months unit of grouping; matrix of order* 7×7. Clearly there are several ways in which a large order matrix may be condensed into one of a smaller order. The method which was used in the present instance was to construct the first row of the condensed canonical form B by interpolating in the integral curve of the original $L_x m_x$ column (1 month units of grouping) for the ages 4·5, 7·5, 10·5, etc., and taking the first differences of the seven values thus obtained. Since interpolation was not very satisfactory in the earlier part of the integral curve—the differences converged rather slowly in this region—the elements were expressed to only three places of decimals. (Some preliminary transformation of the integral $L_x m_x$ figures might have been better in this case.) The characteristic equation, expanded in terms of λ, was found to be

$$\lambda^7 - 1{\cdot}756\lambda^6 - 6{\cdot}899\lambda^5 - 7{\cdot}203\lambda^4 - 5{\cdot}344\lambda^3 - 3{\cdot}244\lambda^2 - 1{\cdot}110\lambda - 0{\cdot}102 = 0.$$

It will be seen that the sum of the coefficients, $R_0 = 25{\cdot}658$, which is necessarily the same as the original net reproduction rate owing to the way in which the coefficients were derived. For interest, the seven roots of this equation were then determined by the root-squaring

method, using 4-place tables of logarithms and Barlow's tables of squares (Whittaker & Robinson, 1932, p. 110). The approximate values of the roots, arranged in descending order of their moduli, are

$$\begin{aligned} \lambda_1 &= 4{\cdot}016, \\ \lambda_2 &= -1{\cdot}032, \\ \lambda_3\lambda_4 &= -0{\cdot}0215 \pm 0{\cdot}6762i \quad (\text{mod.} \doteq 0{\cdot}6765), \\ \lambda_5\lambda_6 &= -0{\cdot}5245 \pm 0{\cdot}3486i \quad (\text{mod.} = 0{\cdot}6298), \\ \lambda_7 &= -0{\cdot}135. \end{aligned}$$

Thus, apart from the positive real root, there are two negative and two pairs of complex roots. It is interesting to note in passing that in this example the modulus of the second latent root is > 1 (vide § 13). From the value of the dominant root we find $r = 0{\cdot}4634$ per head per month of 30 days, an estimate of the rate of increase which is 1·71 % per month higher than that from the large order matrix.

Table 3. *Stable age distributions*

Age group (in months)	$100{,}000b\int_x^{x+3} e^{-rx}l_x\,dx$	From condensed matrix	Summation of 'true' distribution
0–	75,960	75,762	74,452
3–	18,055	18,267	18,905
6–	4,514	4,519	4,939
9–	1,120	1,111	1,280
12–	273	267	326
15–	64	61	80
18–	14	13	18
Total	100,000	100,000	100,000

The net reproduction rate given by the new $L_x m_x$ column which was obtained by working in units of three months, was 25·6162, a figure somewhat lower than the original one of 25·6579. The rate of increase, estimated in a similar way to the former example for one month age units, was $r = 0{\cdot}46034$, again a higher figure, though one of much the same order as that obtained from the condensed matrix.

The appropriate age distributions are given in Table 3, together with the 'true' distribution of Table 2 summed in three month age groups.

Compared with the last column, both of the stable distributions for the data grouped in three month age intervals are tilted towards the younger age classes, so that the number of immature females (< 3 months of age) is overestimated, while the remaining age groups are underestimated. The distributions derived from the integral and from the matrix are again of much the same order, and the differences between them and the last column, although not very great, are quite marked.

The four estimates of the inherent rate of increase which have obtained from these numerical data may be compared in the following table.

In both cases the estimates from the $L_x m_x$ column and from the matrix agree very well: for a given unit of grouping both methods would seem to give comparable results. The differences between the estimates made by the same method are much greater, and the effect

of increasing the unit of grouping, and in this way shortening the labour, is to increase the value of r quite appreciably. Whether, or not, we should regard these estimates for the three months age grouping as satisfactory would depend on the degree of accuracy required in any particular calculation. It must be remembered, however, that the basic numerical data are of rather an extreme type in this example. It is doubtful whether any naturally living rat population would have so good a life table and so high a degree of fertility as that assumed for this imaginary population. In fact it was for these very reasons that these data were

	$L_x m_x$ column	Matrix	Difference
1 month age groups 3 „	0·44565 0·46034	0·44626 0·4634	0·00061 0·0031
Difference	0·01469	0·0171	

chosen as the basis of the numerical calculations in this work. For, if it could be shown that the two methods of computation gave comparable results in this case, it was felt that an even better agreement should be obtained in less extreme examples and more particularly with data relating to populations whose rate of increase is nearer to the stationary state. Although in this example the larger unit of grouping leads to rather unsatisfactory estimates of the rate of increase and the stable age distribution, it seems probable that, for the reasons just given, the differences would be less for instance in the case of human data. Hence, the question of the unit to be adopted is likely to become of less importance in the type of data more commonly met with, though it would be necessary to work out an example for such a population in order to check this point.

There is, however, one way to avoid this difficulty of the working unit for populations with a high relative rate of increase. For example, returning to the numerical data used here, supposing that it was necessary in some particular problem to have a fairly high degree of accuracy in the results, but that the work involved in manipulating the large order matrix of 21×21 was too excessive. It might be sufficient in the case we are imagining to know the age distribution of the population in three month age groups at some particular time in the future, which we will take to be a multiple of three. Then, once the real latent root (λ_1) of the large matrix and its associated stable vector have been determined, it is possible to construct a small order matrix of 7×7 which has λ_1^3 as its dominant root and therefore the same real stable age distribution as the larger matrix, only expressed in three month instead of in one month age units. It is convenient to carry out the calculation in terms of the canonical form B and of ψ vectors. Having determined the dominant latent root and the stable vector for the larger matrix, the elements of the first three rows of B^3 are then written down and summed in columns. This can be done very quickly in the present example, where reproduction does not start until the age of 3 months, for the third row of B^3 is the same as the first row of B; the second row is merely the first row of B shifted one age group to the left; and similarly again for the first row. The sums of the columns are then weighted with the number alive in the appropriate age group in the stable population (ψ_1 vector), and by summing the weighted column totals in groups of three and taking the weighted mean, we obtain the elements of the first row of a 7×7 matrix which has λ_1^3 as its domi-

nant latent root. Thus, the characteristic equation of the original matrix condensed in this way was

$$\lambda^7 - 1{\cdot}5056\lambda^6 - 6{\cdot}4694\lambda^5 - 7{\cdot}2047\lambda^4 - 5{\cdot}5371\lambda^3 - 3{\cdot}4537\lambda^2 - 1{\cdot}3451\lambda - 0{\cdot}1447 = 0,$$

and, out of idle curiosity, all the seven roots were extracted in order to compare them with those of the previous example of a condensed matrix. The estimation in this case was carried out to a higher degree of accuracy. The results were:

$$\begin{aligned}
\lambda_1 &= 3{\cdot}81452, \\
\lambda_2 &= -1{\cdot}02526, \\
\lambda_3\lambda_4 &= -0{\cdot}5905 \pm 0{\cdot}3782i \quad (\text{mod.} = 0{\cdot}70125), \\
\lambda_5\lambda_6 &= 0{\cdot}0280 \pm 0{\cdot}6879i \quad (\text{mod.} = 0{\cdot}68847), \\
\lambda_7 &= -0{\cdot}15876.
\end{aligned}$$

The dominant root of the original matrix was 1·56246 and the cube root of λ_1 is 1·56248. The remaining roots may be compared with those given for the previous example. The two negative roots are very similar and, in the second case, the two pairs of complex roots appear to have changed places, the real part of one pair becoming positive instead of negative. Although the cube root of λ_1 is equal to the dominant root of the original matrix, it is unfortunately not true that a similar relationship holds for the remaining six values of λ. There is, for instance, no negative latent root > 1 for the larger form in this actual example.

This point, however, raises an extremely interesting question. For a given series of data a finite matrix of a relatively small order may be constructed, as in the first example given here of a condensed matrix. Supposing that the order of this matrix is increased step by step and that in each case the latent roots are found. Then, in this approach to an infinite matrix, how do the latent roots behave and what relation does the array of roots in each case bear to those of the preceding steps? For the purposes of comparison it will be necessary to express the roots in terms of some suitable unit of time, e.g. per month or per year. So far as the real positive root is concerned, it seems likely that the series of individual roots will approach nearer and nearer to a limiting value. For the root λ_1 is the ratio N_{t+h}/N_t, or the number of times the stable population has increased at the end of the interval of time h. Then, expressing λ_1 in the chosen unit of time, we have $\Lambda_1 = (\lambda_1)^{1/h}$, and taking logarithms;

$$\log_e \Lambda_1 = \frac{\log_e N_{t+h} - \log_e N_t}{h},$$

so that, when the interval of time becomes very small, corresponding to a matrix of a very large order, and $h \to 0$, the right-hand side of this equation approaches the limit $\dfrac{1}{N}\dfrac{dN}{dt} = \rho$, the true instantaneous relative rate of increase of the stable population. This argument is put forward with a certain amount of diffidence; it is only too easy for the biologist to overlook some flaw which will be immediately obvious to the trained mathematician. But, even if it were a valid argument for the behaviour of the dominant root, it can hardly be extended in this form to the case of the remaining roots; and thus the main question is left unanswered. From the point of view of the biologist, it would be interesting to know whether with an increase in the size of the matrix the array of secondary roots tends to coalesce round certain values of $\lambda^{1/h}$.

16. FURTHER PRACTICAL APPLICATIONS

If we wish merely to estimate the inherent rate of increase and the stable age distribution appropriate to some system of age specific fertility and mortality rates, there is evidently little to choose between the matrix and the ordinary methods of computation. The advantages of expressing the basic rates in the form of a matrix are more clearly seen in considering the type of problem such as the following. Let us suppose that a species of mammal at a certain season of the year invades a fresh environment where there is an ample food supply, a freedom from predators, and plenty of space to accommodate any rapid increase in numbers which might take place. Under these conditions it might be assumed for theoretical purposes that some age specific rates of fertility and mortality would remain approximately constant over a period of time. The age distribution of these immigrants then becomes of some importance owing to the effect which it must necessarily have on the future course of events. For this initial distribution must clearly be very different from that which would ultimately be established in the case of a species, such as a rodent, with possibly a very rapid rate of increase, since nestlings will not be represented in it and young individuals may be present in only relatively small numbers. Supposing then that we have a number of such populations subject to the same age schedules of fertility and mortality, but differing in the age distribution of the original immigrants, we may have reason to enquire how far the development of these populations is affected over a limited period of time by the varying form of this initial distribution, assuming for simplicity that no further waves of immigration occur.

If an estimate of the number and age distribution of the female population at successive intervals of time is alone required, the answer for any form of initial distribution is readily obtained once the series of matrices M, M^2, M^3, ..., M^t have been constructed. But, in addition, we may require to know the changes which might be expected to occur in the birth rate and death rate, and also, for example, in some such rate as the percentage of adult females pregnant, a figure which is one of the simplest measures we have of the degree of fertility among wild populations. Again, in a species like the wild rat we never know the exact age of individuals caught in the field, and thus the only measure of the form of the female age distribution is the percentage of immature females, provided, of course, we are sampling the complete population. Some method is therefore required for calculating such rates at successive intervals of time.

Once the age distribution of the female population at time t is known, an estimate of the expected number of female births per unit of time may be obtained by operating on the age distribution with the maternal frequency figures. Thus, in matrix notation we may write, the number of female births equals $m_x M^t \xi_0$, where ξ_0 is the initial age distribution and the m_x figures are treated as a row vector. Similarly the estimated number of deaths per unit of time may be obtained with the help of the age specific death rates (D_x). The relative rate of increase calculated in this way is not necessarily exact, but it may be sufficiently accurate for our present purposes. As an example of the degree of error involved in this method, we may compare the values of the stable birth rate and death rate, as given in the appendix, with those derived from the matrix stable distribution in Table 2 by operating with the m_x and D_x figures. The latter were in this case computed from the stationary age distribution and the d_x column of the life table ($D_x = d_x/L_x$). The results were as follows:

	'True' values from appendix	By operating with m_x and D_x on matrix stable distribution
Birth-rate (b)	0·51265	0·51257
Death-rate (d)	0·06700	0·06154
$b-d=r$	0·44565	0·45103

The rate of increase is overestimated by about 5·4 per thousand per month, the principal error being in the death rate. This discrepancy is due to the fact that the number of deaths under 2 months of age is underestimated by applying age specific death rates, which are based on the stationary age distribution, to the stable population grouped in one month intervals at these ages. The difference between these distributions happens to be quite marked in this example. The degree of error, however, is not very great; and in the type of problem we are considering, when the age distribution of a population may take any form, this seems to be the only practical method of estimating the rate of increase.*

Supposing, then that in the case of the rat population used here as a numerical illustration, we wished to estimate the number of females, the birth rate and death rate, and the percentage of immature females at monthly intervals up to—say—7 months from the origin of the time scale, when the initial immigration is assumed to take place. Since the jth column of the matrix gives the age distribution of the survivors and the surviving descendants per individual female alive in the age group $j-1$ to j at $t=0$, the sum of the elements in this column gives the number of times the original population in this age group has increased, or decreased, at time t. The percentage of immature females may be obtained once the sum of the first three elements in the column is known (reproduction begins at the age of 3 months in this example); and the number of births and deaths per unit of time may be found by operating on the column with the m_x and D_x figures. Each of these totals, of course, will have to be multiplied in the end by the number of females alive in this age group at $t=0$. Since the initial age distribution may be of any form under the conditions of the problem, it will be necessary first of all to calculate these four totals for every column of each of the seven matrices M^t. Now, to add up the elements forming each column of a matrix is equivalent to premultiplying the matrix by a row vector of units; the sum of the first three elements may be obtained by premultiplying with a row vector of which the first three elements are units and the remainder zeros; and similarly the numbers of births and deaths are found with the help of the row vectors m_x and D_x. Thus, the operations which it is necessary to perform on

* Another similar method, which avoids the actual calculation of the number of deaths by means of the age specific death rates, is suggested by the following relationship. If the transformed age distribution $\psi = H\xi$, where H is the matrix defined in § 5, is operated on with a row vector which consists of the $L_x m_x$ figures (Appendix, Table 4), and the resulting scalar is divided by the sum of the elements of ψ, an estimate is obtained of the relative rate of increase of a population with an age distribution ξ in the original co-ordinate system. This follows from the properties of the transformed population discussed at the end of § 5 and from the relationship between the first row of the canonical form $B = HAH^{-1}$ and the $L_x m_x$ column (§ 6). In the transformed population the death rate $=0$, and the maternal frequency is given by $L_x m_x$. Thus, by transforming back again to the original co-ordinate system

$$r = \frac{[L_x m_x]\, H\xi}{[1]\, H\xi},$$

where [1] defines a row vector of units. Taking ξ as the matrix stable distribution of Table 2, and calculating the row vectors $[L_x m_x]\, H$ and $[1]\, H$, the value of r was estimated by this method to be 0·44468.

each of the columns of M^t may be written as the matrix R, which will consist of $m+1$ columns and n rows, the number of the latter depending on the number of operations. Then the required totals for the matrix M^t will be given by

$$Z^t = RM^t = RMMM \dots M,$$

and it is easy to see how the Z matrices may be built up in succession without calculating the actual matrices M^2, M^3, M^4, etc. Once the series of Z matrices have been constructed, we can obtain from $Z^t\xi_0$ the necessary figures from which the required rates at time t for a population with an initial age distribution ξ_0 may be calculated. Moreover, if we wish, the contributions made, for instance, to the total number of births or deaths by any particular age group in the initial distribution can also be determined.

The computations in this illustration have been greatly simplified by the assumption that the system of age specific fertility and mortality rates remains constant. In the case when the basic matrix M is changing with time and the age distribution at time t is given by $M_t \dots M_3 M_2 M_1 \xi_0$, some of the rows of R will also be varying. Hence the series of Z matrices could not be built up without first computing $M_2 M_1$, $M_3 M_2 M_1$, etc. The latter, however, may often be of interest in themselves. For, if each column of M^t—or $M_t \dots M_3 M_2 M_1$ in the case of a variable matrix—is multiplied by the number of females alive in the appropriate age group at $t = 0$, the complete age structure of the population at time t is represented in the form of a two-dimensional array. Since the sum of the elements in each row is the total number alive in the age group x to $x+1$ at time t, the number contributed to this total by each age group at $t = 0$ is given by the individual entries.

APPENDIX

(1) *The tables of mortality and fertility*

The basic life table and fertility table which have been used in the numerical part of this study are given in Table 4. The adult l_x figures from the age of 2 months onwards are based on the mortality observed among the females of a domesticated brown rat stock housed at the Wistar Institute, Philadelphia. According to the data for 26 generations of this laboratory stock published by King (1939) it appears that out of 1384 females alive at the age of 2 months (60 days), 1337 were alive at 12 months, and 984 at 20 months. This information gave three points on the l_x curve, supposing that these survivors could be regarded as ordinates at these exact ages. In order to interpolate for other ages, a logistic type of curve was fitted to the data, the values of the constants being chosen so that the curve passed through these three points. The l_x values in Table 4 are given by

$$l_x = \frac{0{\cdot}85156355}{1 + 0{\cdot}00101065 e^{0{\cdot}30016x}}, \quad \text{for} \quad x \geqq 2.$$

Although the original data did not extend beyond the age of 20 months, by which time the vast majority of the females had ceased breeding (King, 1939), this l_x curve was extrapolated to later ages, whenever necessary, simply for the purposes of this theoretical study.

The degree of infant mortality assumed here, namely 15 % between birth and the age of 2 months, is entirely arbitrary; it represents a moderate degree of loss at these early ages. Some care, however, was taken to weld the infant mortality smoothly on to the remainder of the l_x curve, and it was assumed that the number of deaths according to age (d_x) decreased

geometrically between birth and the age of 2 months. The actual calculations for these age groups were carried out in units of 1/8 of a month and the resulting l_x curve was integrated by means of Simpson's rule. The same method of numerical integration was used also for the adult part of the life table in order to obtain the L_x figures.

The fertility table is partly artificial and was constructed in the following way. The gross reproduction rate of these domesticated brown rats was estimated from the data published by King (1939) to have been just under 10 litters for the later generations, when the stock was thoroughly adapted to life in the laboratory. The frequency of litter production according to the age of the mother has been found by the author (unpublished observations) to be represented closely by a Pearsonian type I curve in the case of certain litter fertility tables, for example in the cross-albino rat, the vole and some human populations with a high degree of fertility; and, moreover, the values of β_1 and β_2 were very similar for all three species. The actual equation for the curve used here may be written

$$y = y'x^{1\cdot5-1}(a-x)^{2\cdot5-1},$$

and the range was assumed to be from 3 to 21 months, which for grouping purposes represents the span of the reproductive ages observed by King in this Wistar strain of brown rats. The ordinates of the integral curve of the above equation were taken from the tables of the incomplete Beta function and, in this way, a column was formed which gave a gross reproduction rate of 10 litters. The individual entries were then multiplied by the mean number of daughters per litter according to the age of the mother, which was recorded by King, and thus the m_x figures in Table 4 were obtained. The gross reproduction rate is 31·21 daughters and the net rate 25·66. These tables of fertility and mortality were originally constructed in order to determine the relative rate of increase and the type of stable age distribution which might be expected in a brown rat population living under more or less optimum conditions.

(2) *Calculation of the rate of increase*

Some difficulty was experienced in obtaining a satisfactory estimate of the rate of increase (r) from the usual solution (Dublin & Lotka, 1925) of the equation:

$$\int_0^\infty e^{-rx}l_x m_x dx = 1.$$

The 4th degree equation in r with the numerical coefficients based on the seminvariants of the $L_x m_x$ distribution, the estimates of which are given in § 6, was

$$4\cdot90199r^4+3\cdot69283r^3-7\cdot07198r^2+9\cdot60604r-3\cdot2448498 = 0,$$

and the real root was found to be 0·42447. This value of r was, however, clearly too low and a better estimate had to be obtained in a rather roundabout way, since it was thought that the use of higher moments than the fourth would be unsatisfactory in the present example.

If the force of mortality represented by the original life table is increased by a constant factor (r) which is independent of age, the new life table is

$$l'_x = e^{-rx}l_x,$$

and the net reproduction rate will be given by $R_r = \Sigma L'_x m_x$, where L'_x are the integrals of the new l'_x curve. Clearly, the greater r is taken to be, the smaller R_r becomes. Then, suppose that the relation between R_r and r is given by

$$\log_e R_r = a+br+cr^2+dr^3+\ldots,$$

Table 4

Age (x) Units of 30 days	Life table		Fertility table	
	l_x	L_x	m_x	$L_x m_x$
0	1·00000	0·46544	—	—
0·5	0·88706	0·43489	—	—
1	0·85882	0·42725	—	—
1·5	0·85176	0·42534	—	—
2	0·85000	0·84973	—	—
3	0·84945	0·84910	1·1342	0·96305
4	0·84871	0·84824	2·0797	1·76408
5	0·84772	0·84708	2·6596	2·25289
6	0·84638	0·84553	2·8690	2·42582
7	0·84458	0·84344	2·9692	2·50434
8	0·84217	0·84063	2·9535	2·48280
9	0·83893	0·83687	2·8143	2·35520
10	0·83459	0·83184	2·6114	2·17227
11	0·82881	0·82515	2·2455	1·85287
12	0·82113	0·81629	2·0533	1·67609
13	0·81098	0·80463	1·7971	1·44600
14	0·79768	0·78940	1·5561	1·22839
15	0·78039	0·76975	1·2175	0·93717
16	0·75821	0·74472	0·9548	0·71106
17	0·73018	0·71342	0·6610	0·47157
18	0·69548	0·67512	0·4043	0·27295
19	0·65354	0·62953	0·1846	0·11621
20	0·60434	0·57696	0·0435	0·02510
21	0·54859	—	—	—
Total	—	—	31·2086	25·65786

where $a = \log_e R_0$, and the constants b, c, d, etc. are to be determined. It was assumed in the present instance that a 4th degree polynomial in r would be sufficient, and four new life tables were constructed taking r to be in turn 0·1, 0·2, 0·4 and 0·5. The L'_x integrals were obtained by Simpson's rule for the reproductive ages and the four values of R_r calculated. The equations for finding the values of the constants were:

$$0{\cdot}0001e + 0{\cdot}001d + 0{\cdot}01c + 0{\cdot}1b = -0{\cdot}8934974,$$
$$0{\cdot}0016e + 0{\cdot}008d + 0{\cdot}04c + 0{\cdot}2b = -1{\cdot}6708610,$$
$$0{\cdot}0256e + 0{\cdot}064d + 0{\cdot}16c + 0{\cdot}4b = -2{\cdot}9792984,$$
$$0{\cdot}0625e + 0{\cdot}125d + 0{\cdot}25c + 0{\cdot}5b = -3{\cdot}5490542,$$

whence

$$b = -9{\cdot}617235,$$
$$c = 7{\cdot}371816,$$
$$d = -5{\cdot}698291,$$
$$e = 2{\cdot}062332.$$

Inserting these values of the constants in the equation for $\log_e R_r$, the value of r for which $\log_e R_r = 0$ was found to be 0·44565.

The stable birth rate was estimated in the usual way from

$$\frac{1}{b} = \int_0^\infty e^{-rx} l_x \, dx.$$

The integrals were computed by means of Simpson's rule, treating the age groups 0–2 separately from the remainder of the life table, the units adopted being 1/4 month for the early ages compared with 1 month for the later. It was found that

$$\int_0^{21} e^{-0\cdot 44565x} l_x dx = 1\cdot 95064,$$
$$b = 0\cdot 51265,$$

and hence the death rate,

$$d = 0\cdot 06700.$$

The value of r is so high in this case that the error in the estimate of b due to neglecting the ages from 21 onwards would only be in the last figure. The stable age distribution is given in Tables 1 and 2 of the text. Owing to the way in which the value of r was determined, it will be found that the birth rate of the stable population obtained by operating on this age distribution with the maternal frequency figures is precisely the same as that given by the above integral.

(3) *Numerical values of the matrix elements*

The numerical elements of the matrices A and B to which reference has been made in §§ 3 and 6 of the text are given in Table 5.

Table 5

Age group x	Matrix A		Matrix B
	P_x	F_x	Elements of first row
0–	0·94697	0	0
1–	0·99665	0	0
2–	0·99926	0·3964	0·3741
3–	0·99899	1·4939	1·4089
4–	0·99863	2·1777	2·0517
5–	0·99817	2·5250	2·3756
6–	0·99753	2·6282	2·4682
7–	0·99667	2·6749	2·5059
8–	0·99553	2·6018	2·4293
9–	0·99399	2·4419	2·2698
10–	0·99196	2·1865	2·0202
11–	0·98926	1·9044	1·7454
12–	0·98572	1·7259	1·5648
13–	0·98107	1·4918	1·3332
14–	0·97511	1·2415	1·0885
15–	0·96748	0·9522	0·8141
16–	0·95797	0·7141	0·5907
17–	0·94631	0·4618	0·3659
18–	0·93247	0·2518	0·1888
19–	0·91649	0·0901	0·0630
20–		0·0035	0·0022

I wish to thank Professor M. Greenwood, F.R.S. and Dr J. O. Irwin for their kindness in reading the manuscript, and also Dr H. Motz for many fruitful discussions. The study arose out of some research work which is being carried out by the Bureau of Animal Population with the aid of a grant from the Agricultural Research Council.

REFERENCES

Aitken, A. C. (1939). *Determinants and Matrices*, 1st ed. Edinburgh: Oliver and Boyd.

Bernardelli, H. (1941). Population waves. *J. Burma Res. Soc.* **31**, no. 1, 1–18.

Chapman, R. N. (1933). The causes of fluctuations of populations of insects. *Proc. Hawaii Ent. Soc.* **8**, no. 2, 279–92.

Charles, Enid (1938). The effect of present trends in fertility and mortality upon the future population of Great Britain and upon its age composition. Chap. II in *Political Arithmetic*, ed. Lancelot Hogben. London: Allen and Unwin.

Dublin, L. I. & Lotka, A. J. (1925). On the true rate of natural increase as exemplified by the population of the United States, 1920. *J. Amer. Statist. Ass.* **20**, 305–39.

Frazer, R. A., Duncan, W. J. & Collar, A. R. (1938). *Elementary Matrices*. Camb. Univ. Press.

Glass, D. V. (1940). *Population Policies and Movements in Europe*. Oxford: Clarendon Press.

King, Helen D. (1939). Life processes in grey Norway rats during fourteen years in captivity. *Amer. Anat. Mem.* no. 17, 1–72.

Turnbull, H. W. & Aitken, A. C. (1932). *An Introduction to the Theory of Canonical Matrices*. London and Glasgow: Blackie and Son.

Whittaker, E. T. & Robinson, G. (1932). *The Calculus of Observations*, 2nd ed. London and Glasgow: Blackie and Son.

ERRORS IN THE ROUTINE DAILY MEASUREMENT OF THE PUERPERAL UTERUS

By C. SCOTT RUSSELL, M.A., M.B., Ch.B., F.R.C.S.Ed., M.R.C.O.G.
First Assistant, Nuffield Department of Obstetrics and Gynaecology, University of Oxford

In both district and institutional midwifery, it is usual to measure each day the height of the puerperal uterus above the symphysis pubis. Though most observers would say that the purpose behind this custom is the detection of uterine subinvolution, there is no agreement as to the precise conditions that must be fulfilled or the criteria required for establishing such a diagnosis. The present paper is concerned with the systematic and the uncontrolled errors of this routine measurement, and with the significance of fluctuations in the observed height of the puerperal uterus.

In institutions, the puerperal uterus is usually measured by whichever member of the nursing staff happens to be available: often the measurements are made by pupil midwives. It must be rare for a uterus in any period of 7 days to be measured by only one observer. The measurement, which is made between the upper surface of the uterus and the upper surface of the symphysis pubis, is recorded on the temperature chart; it is probably never recorded more accurately than to the nearest ¼ in., and usually the nearest ½ in. is considered sufficiently accurate.

Normal uterine involution

It is widely taught that the normal human uterus involutes ½ in. a day during the puerperium. This is only an approximation; if it were true, the puerperal uterus having shrunk in 1 day the last ½ in., would suddenly stop involuting. This would be contrary to clinical experience which tells us that the process starts fast and finishes slowly. The changing daily amount of normal involution would be too cumbersome a standard in practice; besides, great accuracy is not necessary when the uterus is measured only once a day and slight irregularities are of little clinical importance. A rate of ½ in. a day may therefore be taken as accurate enough for practical purposes.

Factors influencing the recorded height of the uterus

The factors that influence the recorded height of the puerperal uterus may be divided into two broad groups. First, there are those which act constantly in one direction, for example, the distended bladder or lower bowel. Secondly, there are the uncontrolled factors, such as the thickness of the abdominal wall, the mobility of the uterus, the uterine tone, the method of measurement, the points chosen for measuring, the individual differences between observers and so on. Though these many factors are just as likely to cause the reading to be too high as too low, they are important because together they constitute the error of the measurement; without knowledge of this error, logical deductions based on the measurements are not possible. The following five studies were made to obtain estimates of this error.

Method

First investigation

Seven puerperal patients whose uteri could be felt were chosen, and six unpractised observers measured each uterus twice, the interval between measurements being about

15 min. For convenience, medical students were used as observers; it was thought that they would be neither more nor less accurate than pupil midwives. An effort was made to prevent conscious or subconscious bias on the part of each observer (1) by explaining the nature of

Table 1. *Duplicate measurements of puerperal uterus by unpractised observers*

Observers' duplicate readings	Patients 1	2	3	4	5	6	7	Observers' totals	
1	5·00	5·00	4·25	4·00	4·25	5·75	4·75	33·00	68·25
	5·50	4·25	4·75	4·50	4·75	6·00	5·50	35·25	
2	4·75	4·75	4·25	4·75	4·50	5·50	5·00	33·50	65·00
	5·00	4·25	3·75	4·00	4·25	5·25	5·00	31·50	
3	5·50	4·75	4·00	4·25	4·75	5·75	5·00	34·00	69·50
	5·75	4·75	4·50	4·50	4·75	5·75	5·50	35·50	
4	6·00	5·00	4·50	3·75	5·50	5·50	5·50	35·75	76·25
	6·75	5·75	5·25	4·50	6·50	6·00	5·75	40·50	
5	5·25	4·50	4·00	4·00	5·75	6·00	4·50	34·00	70·25
	5·75	4·75	5·00	4·50	5·25	5·75	5·25	36·25	
6	5·25	4·50	4·25	4·25	6·00	5·50	5·25	35·00	72·00
	6·00	5·00	5·00	4·75	5·75	5·75	4·75	37·00	
Patients' totals	31·75	28·50	25·25	25·00	30·75	34·00	30·00	205·25	
	34·75	28·75	28·25	26·75	31·25	34·50	31·75	216·00	
	66·50	57·25	53·50	51·75	62·00	68·50	61·75		421·25

Analysis of variance

	Sum of squares	Degrees of freedom	Mean square	Variance ratio	Probability
Main effects					
Patients	19·9349	6	3·3225	43·6310	Less than 0·001
Observers	5·0845	5	1·0169	13·3539	Less than 0·001
Readings	1·3778	1	1·3778	18·0932	Less than 0·001
First order interactions					
Patients-observers	5·6187	30	0·1873	2·4576	0·05–0·01
Patients-readings	0·6798	6	0·1133	1·4879	Greater than 0·20
Observers-readings	1·6891	5	0·3378	4·4360	Less than 0·01
Second order interactions					
Patients-observers-readings [error]	2·2845	30	0·0762		
Total	36·6693	83			

the investigation, and (2) by ensuring that previous readings of other observers were not known until subsequent readings were recorded. Measurements to the nearest $\frac{1}{4}$ in. were requested. In Table 1 the measurements are recorded together with the corresponding analysis of variance. This analysis and certain important facts arising from it must be considered in detail.

Analysis of variance

The total variability of the eighty-four separate observations may be expressed by a 'mean square', i.e. the sum of squares of deviations of each observation from the arithmetic mean divided by the degrees of freedom. This variability, as expressed by a 'mean square', is divisible into mean squares due to the main effects and the first order interactions together with a remainder which is the best first estimate of uncontrolled error. Based on 30 degrees of freedom this has a mean square of 0·076, against which all the other mean squares are compared to give the different variance ratios. From these the probabilities are found from the appropriate tables (Fisher & Yates, 1943).

The interaction 'patients-readings' as expressed by the corresponding variance ratio is not significant;* the interval between readings, therefore, appeared to have had a similar effect on the seven patients. The interaction 'observers-readings' is highly significant.* This is found, by an analysis of variance on the first and second readings taken separately, to be entirely due to the second readings. Whereas the variance ratio 'between observers' obtained from the first readings is not significant (0·933), that from the second readings is highly significant (11·03). A possible explanation is that the observers largely responsible subconsciously altered the technique of measurement between reading 1 and reading 2. The interaction 'patients-observers' is significant,* suggesting that the observers found the different uteri difficult to measure in different ways.

Passing to the main effects, the variance ratio 'patients' is highly significant, as would be expected from patients observed on different days of the puerperium. The variance ratio 'observers' is highly significant, emphasizing the importance of the personal factor. The variance ratio 'readings' is also highly significant; this may be due to an accumulation of urine between the first and second readings causing slight upward displacement of the uterus.

The analysis of variance has been taken to the limit. Of the three first order interactions, the first and third are not controllable with unpractised observers; the second is not significant. All three may therefore be included in the estimate of uncontrolled error. The mean square of the adjusted error is now 0·1447. The variance of the difference between two observations is therefore 0·2894 and the standard error of the difference, the square root of this, i.e. 0·5380. If the effect of different observers is taken into account, the mean square 'observers' must be included in the error estimate. The mean square for error is then 0·2021, and the standard error of the difference between two readings 0·6358.

Second investigation

A similar study was carried out with seven other unpractised observers and seven other patients. In addition to the duplicate readings being separated by an interval of time—approximately 15 min.—each patients was given 0·5 mg. ergometrine by mouth as soon as the first readings had been completed. The results are given in Table 2. All the main effects and all the first order interactions are highly significant, supporting the evidence already presented that the measurement not only is difficult but also is influenced by many factors. The three first order interactions would not be controlled in practice, so that all may reasonably be included in the error estimate. This addition alters the mean square for error to 0·1621. The standard error of the difference between two observations is therefore 0·5694,

* The term 'significant' is used where the probability is between 0·05 and 0·01. The term 'highly significant' is used where the probability is less than 0·01. 'Not significant' refers to a probability greater than 0·05.

which is in close agreement with the value 0·5380 obtained from the first investigation. The inclusion of the effect 'between observers' raises the standard error of the difference to 0·6821, which also agrees closely with the corresponding figure 0·6358 obtained from the first investigation.

Table 2. *Duplicate measurements by unpractised observers before and after ergometrine.*

Observers' readings before and after ergometrine	Patients 1	2	3	4	5	6	7	Observers' totals	
1	5·25 4·75	5·50 4·50	3·25 2·75	4·75 3·50	6·25 5·50	6·50 6·00	4·50 5·00	36·00 32·00	68·00
2	6·00 5·25	5·50 5·50	3·00 3·00	4·25 4·50	6·50 6·00	6·50 6·00	5·00 5·25	36·75 35·50	72·25
3	4·50 4·50	4·75 4·50	3·00 2·75	4·50 4·25	6·00 5·50	5·75 4·50	4·75 4·25	33·25 30·25	63·50
4	4·50 4·50	5·50 5·00	4·25 3·00	5·00 4·25	5·75 5·00	5·50 4·50	4·75 4·50	35·25 30·75	66·00
5	5·00 5·00	5·50 5·50	2·75 3·00	4·75 5·00	6·00 6·50	6·50 6·00	5·50 5·75	36·00 36·75	72·75
6	4·50 4·75	5·25 5·00	2·75 2·25	4·50 4·75	6·00 5·75	5·75 4·75	4·25 4·50	33·00 31·75	64·75
7	5·50 5·75	5·25 5·75	3·50 3·50	4·00 5·00	6·25 6·25	6·25 5·75	5·25 5·50	36·00 37·50	73·50
Patients' totals	35·25 34·50	37·25 35·75	22·50 20·25	31·75 31·25	42·75 40·50	42·75 37·50	34·00 34·75	246·25 234·50	
	69·75	73·00	42·75	63·00	83·25	80·25	68·75		480·75

Analysis of variance

	Sum of squares	Degrees of freedom	Mean square	Variance ratio	Probability
Between patients	76·4707	6	12·7451	182·0707	Less than 0·001
Between observers	7·3190	6	1·2198	17·4257	Less than 0·001
Between readings	1·4089	1	1·4089	20·1271	Less than 0·001
Patients-observers	7·3061	36	0·2029	2·8986	Less than 0·01
Patients-readings	1·5420	6	0·2570	3·6714	Less than 0·01
Observers-readings	2·2473	6	0·3745	5·3500	Less than 0·001
Error	2·5205	36	0·0700		
Totals	98·8145	97			

Error of practised observers (third investigation)

The error of measurement with unpractised observers was of special importance for those institutions where the routine puerperal uterine measurements are made by pupil midwives. Elsewhere the routine is the responsibility of comparatively senior nurses. In order to

obtain an estimate of the error of trained observers another simple study was made. Ten puerperal patients each had the uterus measured by five observers of whom four were nurses and one was a doctor. The observers were experienced in the measurement; they were asked to use the method to which they were accustomed. Each observer measured each uterus once. In Table 3 the measurements are tabulated with the corresponding analysis of variance.

It should be noted that not one of the measurements was given more accurately than to the nearest $\frac{1}{4}$ in., and that forty-four out of the fifty readings were given to the nearest $\frac{1}{2}$ in.

The two main effects, 'between patients' and 'between observers', are both highly significant, as would be expected. The remainder or error estimate has a mean square of 0·3308, which is considerably higher than that of the unpractised observers. This apparent paradox

Table 3. *Measurements of the puerperal uterus by practised observers*

Observers	Patients										Observers' totals
	1	2	3	4	5	6	7	8	9	10	
1	5·00	5·00	5·00	3·00	4·00	4·50	2·50	4·00	4·00	6·25	43·25
2	4·25	4·00	5·75	1·50	2·50	3·00	1·00	2·50	3·00	5·00	32·50
3	4·50	4·50	5·00	2·00	2·50	2·00	3·00	1·50	4·00	6·00	35·00
4	4·50	5·00	6·00	2·50	4·50	3·00	1·50	4·00	3·50	5·50	40·00
5	4·50	5·00	5·50	3·00	4·00	4·25	2·25	3·50	4·25	6·50	42·75
Patients' totals	22·75	23·50	27·25	12·00	17·50	16·75	10·25	15·50	18·75	29·25	193·50

Analysis of variance

	Sum of squares	Degrees of freedom	Mean square	Variance ratio	Probability
Between patients	70·2800	9	7·8088	23·608	Less than 0·001
Between observers	9·0925	4	2·2731	6·87	Less than 0·001
Error	11·9075	36	0·3308		
Total	91·2800	49			

is probably largely due to the fact that the unpractised observers were asked to measure to the nearest $\frac{1}{4}$ in. and thereby encouraged to be accurate, whereas four out of five of the practised observers made the measurement to the nearest $\frac{1}{2}$ in.—except on three occasions—as this was their usual practice.

Error of individual observers (investigations 4 and 5)

Two studies were now made to obtain estimates of the error of individual persons.

An experienced midwifery sister tutor made five replicate readings to the nearest $\frac{1}{4}$ in. on nine puerperal patients. The results are shown in Table 4. The mean square between readings was not significant and was therefore included in the error estimate. The mean square for error was now 0·1639, which is not appreciably different from 0·1447 and 0·1621 which are the corresponding figures previously obtained.

Having had special experience of the measurement, I repeated the above study. Of twelve unselected puerperal patients only eight were measured, the remaining four being thought unsuitable because of abdominal tenderness (one case), recent delivery (one case), and difficulty in palpating the uterus (two cases). This partial selection of cases left for analysis only those whose uteri could be felt fairly easily. The results are shown in Table 5. The mean square between readings was again not significant and was included in the estimate of error. The new value was 0·0813, which is lower than the previous best. The improvement, though partly the result of greater care in measurement, is probably largely due to some selection of material.

Table 4. *Replicate measurements by midwifery sister tutor*

Replicate readings	Patients									Readings' totals
	1	2	3	4	5	6	7	8	9	
1	5·00	5·75	3·00	2·00	3·75	3·00	4·00	4·50	6·25	37·25
2	5·00	6·00	2·75	2·75	3·75	3·00	5·00	4·50	6·00	38·75
3	5·00	5·00	3·00	3·00	4·75	3·00	4·75	4·50	6·00	39·00
4	4·75	6·00	3·50	4·00	4·75	3·25	4·50	4·25	6·00	41·00
5	4·50	5·00	3·00	3·75	4·50	3·00	4·75	4·50	6·00	39·00
Patients' totals	24·25	27·75	15·25	15·50	21·50	15·25	23·00	22·25	30·25	195·00

Analysis of variance

	Sum of squares	Degrees of freedom	Mean square	Variance ratio	Probability
Between patients	47·9750	8	5·9969	37·565	Less than 0·001
Between readings	0·7916	4	0·1979	1·240	Greater than 0·20
Error	5·1084	32	0·1596		
Readings + error	5·9000	36	0·1639		
Totals	53·8750	44			

Effect of the bladder on the puerperal uterus

Attention must now be directed to the factors which bias the recorded height of the uterus. The only factor which constantly makes the recorded size smaller than reality is a tendency for the uterus to fall backward into the pelvis towards the end of the first week of the puerperium. There are, however, two important influences tending to make the uterus seem larger than reality. The one is the urinary bladder, of special importance in the immediate post-partum period: the other is the pelvic colon and rectum, of more importance a few days later. The effect of the bladder is well known, and I have even found the uterine fundus displaced so far upwards by the bladder that it could be felt under the costal margin and ballotted like the foetal head in a breech presentation. To show this upwards displacement the following short study was made.

Table 5. *Replicate measurements by author*

Replicate observations	Patients 1	2	3	4	5	6	7	8	Readings' totals
1	4·75	3·50	2·50	3·00	5·00	5·25	7·00	4·25	35·25
2	4·25	3·25	2·00	3·00	5·25	5·25	6·50	4·75	34·25
3	4·50	3·50	2·50	3·00	5·50	5·25	6·00	4·50	34·75
4	4·50	3·50	2·00	3·25	5·00	4·75	7·00	4·50	34·50
5	3·75	3·25	2·25	3·50	4·75	4·75	6·50	4·75	33·50
Patients' totals	21·75	17·00	11·25	15·75	25·50	25·25	33·00	22·75	172·25

Analysis of variance

	Sum of squares	Degrees of freedom	Mean square	Variance ratio	Probability
Between patients	64·4609	7	9·2087	107·83	Less than 0·001
Between readings	0·2093	4	0·0523		
Patients-readings [error]	2·3907	28	0·0854		
Readings + error	2·6000	32	0·0813		
Totals	67·0609	39			

Fourteen unselected patients were used. Each patient's uterus was measured by myself just before a morning bed-pan round. Following the use of the bed pan each uterus was again measured. In Table 6 the duplicate readings and the differences are tabulated. Being concerned with the displacement of the uterus by the bladder the appropriate statistical test is based on the difference column, in which it should be noted all fourteen figures bear the same sign. Using the t test (Fisher, 1941, p. 117), the probability is found to be less than

Table 6. *Measurements of the puerperal uterus before and after micturition*

Measurements First	Second	Difference [y]	Measurements First	Second	Difference [y]
7·50	6·25	1·25	6·25	5·25	1·00
4·25	3·75	0·50	5·00	2·00	3·00
7·00	4·50	2·50	6·75	5·00	1·75
6·00	4·50	1·50	7·50	5·25	2·25
8·00	4·25	3·75	4·25	2·50	1·75
3·75	2·25	1·50	6·50	4·25	2·25
5·00	4·00	1·00	5·50	4·00	1·50

$$\text{Sum}\,\frac{(y-\bar{y})^2}{13} = 0{\cdot}7445 \qquad \frac{0{\cdot}7445}{14} = 0{\cdot}0532 = (-0{\cdot}231)^2.$$

$$\bar{y} = 1{\cdot}8214 \quad t_{13} = 7{\cdot}8 \quad P < 0{\cdot}001.$$

0·001. As the condition of the bladder was the only important factor that varied between reading 1 and reading 2, there can be no doubt that the uterus had been displaced upwards.

Because one would expect a high correlation between the displacement of the uterus and the amount of urine passed, each difference was compared with the corresponding output of urine. The regression coefficient of fall in uterine height on urinary output is 0·046. The mean square for error is very high—further evidence of the difficulty in measuring the puerperal uterus—and the regression coefficient is, in fact, not significantly different from zero (P between 0·1 and 0·2). If further cases were studied, one would expect from clinical experience that a significant coefficient would be obtained.

Discussion

Though we cannot tell in advance precisely how much a given uterus will involute in a given time, it is known that the average involution rate of the healthy organ is, as a close approximation, ½ in. per day during the early days of the puerperium. Therefore, the expected height of a uterus on any one of these days may be estimated as ½ in. less than that recorded for the previous day.

The records of the uterine height are influenced by many factors outside clinical control; these are of real importance because together they constitute the error of measurement. Without knowledge of the magnitude of this error, it is impossible to judge the significance of measurements and differences between measurements.

Five investigations have been made to determine the magnitude of this error; two were with unpractised observers and three with practised observers. The standard error of the difference between two readings of a single average careful observer is approximately ½ in. Unless, therefore, the difference between the recorded height of the uterus and the expected height is more than 1 in., i.e. twice the standard error of the difference, it is not safe to conclude that anything other than the error of the measurement has been responsible.* But, according to our estimate, the uterus should involute only ½ in. per day. Therefore, unless the recorded height is at least ½ in. greater than that recorded for the previous day, we should not be surprised nor assume anything to be wrong, as such a difference is likely to be due to this error only. This argument may be extended to cover comparisons between measurements at intervals up to 2 or 3 days. If taken much beyond 3 days, however, a serious fallacy might result. Our estimate of the average daily uterine involution is an approximation, and as this is multiplied, so will the error of the approximation increase.

By exercising great care, measuring to the nearest ¼ in., and selecting only those patients whose uterus can be felt easily, it should be possible to reduce the standard error of the difference between two measurements to the region of ¼ in. Any investigation designed to compare involution rates under different conditions would be helped by such increase in accuracy.

There are two common controllable influences which elevate the puerperal uterus. The one, more important in the early days of the puerperium, is a full bladder; the other, exerting its influence later, and of less importance, is a full bowel (Moir & Russell, 1943). It has now been shown, not only that the full bladder raises the uterus, but also that in

* Strictly the comparison between normal and abnormal uterine involution involves a test of significance between two regression lines. In practice, however, it would be unreasonable to expect any medical attendant to calculate the regression line on every occasion that the size of the uterus was larger than expectation. The method described should be adequate and, requiring only mental arithmetic, is suitable for the bedside.

thirteen out of fourteen patients with normal bladder distension the extent was up to or greater than twice the standard error of the difference between two measurements of uterine height. With pathological bladder distension, the upward displacement of the uterus will be still greater and correspondingly more than the error of measurement; in contrast, alteration in uterine height due to subinvolution will be less than the error of measurement.

The practical conclusion is that careful routine daily puerperal uterine measurement is of value, more because it helps in the recognition of abnormal states of the bladder or bowel than those of the uterus. Distension of the bladder or bowel should always be suspected if, from any two measurements over a period of not more than 3 or 4 days, the observed height of the uterus is more than 1 in. greater than the expected height. It must, however, be emphasized that this figure of 1 in. has been obtained from measurements carefully made to the nearest $\frac{1}{4}$ in.; it must be increased to $1\frac{1}{2}$ or even 2 in. if measurements are perfunctorily made by different persons each day. A diagnosis of uterine subinvolution is only permissible after due allowance has been made for the effect of all factors, controllable and uncontrollable, that influence the recorded height.

I wish to record my thanks for help in the preparation of this paper to Mr P. H. Leslie, Dr W. T. Russell and Professor Chassar Moir.

REFERENCES

FISHER, R. A. (1941). *Statistical Methods for Research Workers.* Oliver and Boyd.

FISHER, R. A. & YATES, F. (1943). *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd.

MOIR, J. C. & RUSSELL, C. S. (1943). *J. Obstet. Gynaec.* **50**, 94.

ON A METHOD OF ESTIMATING FREQUENCIES

BY J. B. S. HALDANE, F.R.S.

In a very large variety of investigations we desire to estimate the frequency of an attribute in a series of populations, each of which is so much larger than the sample taken from it that it may be regarded as infinite. If p be the frequency of the attribute, $q = 1-p$, and n the number in the sample, the standard error of the estimate of p is of course $\sqrt{(pq/n)}$. Provided p does not vary much from one sample to another, it is desirable to keep n approximately constant. But when p varies greatly this is unsatisfactory. Thus if $n = 1000$, and $p = 0{\cdot}3$, its standard error is 0·015, but if p is 0·01 its standard error is 0·0031, so that we could not distinguish between populations where p was 0·01 and 0·005. In such a case it may be desired that the standard error of each value of p should be roughly proportional to p rather than to its square root.

Such cases arise in haematology. My friend Dr R. A. M. Case has been investigating the frequency of siderocytes, an abnormal type of red blood corpuscle described by Grüneberg (1942), in a number of bloods. Their frequency ranged from about 0·2 to 20 % or more. He has adopted the method of counting stained red corpuscles in a film until he had counted some definite number m, usually about 20, of siderocytes. If, in order to count this number, he had observed a total of n red corpuscles, he took $\frac{m-1}{n}$ as his estimate of p. It will be shown that the correct estimate is $\frac{m-1}{n-1}$, so that his error was negligible; and the standard error will be calculated.

Let p be the frequency of abnormal cells, and $q = 1-p$,
m be the number of abnormal cells counted,
n be the total number of cells counted,
$x = \frac{m-1}{n-1}$ be the estimate of p.

Clearly n may have any positive value exceeding $m-1$. Let w_n be the probability that exactly n cells are counted before m abnormal cells are observed. Two things are necessary and sufficient if this is to be the case. The first $n-1$ cells must include $m-1$ abnormals, and the nth cell must be abnormal. The probabilities of these two events are $\binom{n-1}{m-1} p^{m-1}q^{n-m}$, and p. Hence $w_n = \binom{n-1}{m-1} p^m q^{n-m}$. This is the coefficient of t^n in $\left(\frac{pt}{1-qt}\right)^m$.

The mean value of x is

$$\bar{x} = (m-1) \sum_{n=m}^{\infty} \frac{w_n}{n-1}$$

$$= \sum_{n=m}^{\infty} \binom{n-2}{m-2} p^m q^{n-m},$$

or if $n-m = r$

$$p^m \sum_{r=0}^{\infty} \binom{m+r-2}{r} q^r$$

$$= p^m(1-q)^{1-m}$$

$$= p.$$

Thus x is an unbiased estimate of p.

The modal value of x is the value given by the value of n which makes w_n maximal. $\frac{w_{n+1}}{w_n} = \frac{nq}{n-m-1}$. Hence w_n exceeds w_{n-1} and w_{n+1}, if n lies between $\frac{m-1}{p}$ and $\frac{m-1}{p}+1$. Therefore the modal value of x ranges between p and $p+\frac{p^2}{m-1-p}$. Thus the distribution of x is slightly asymmetrical, as is obvious from the fact that since n can have any value exceeding $m-1$, the range of possible values of x is from 0 to 1. If we used m/n as our estimate of p a bias would be introduced. In fact, the mean value of m/n can be shown to be

$$mp^m q^{-m}\int_0^q t^{m-1}(1-t)^{-m}\,dt, \quad \text{or} \quad p\sum_{r=0}^{\infty}\frac{m!\,r!\,q^r}{(m-r)!}.$$

On the other hand, the mean value of n/m is easily shown to be p^{-1}. Thus if we were in the habit of expressing frequencies in such a form as '1 in 20' rather than '5 %' or '0·05', the method of counting up to a fixed number m of abnormals, and dividing this into the total number n, would give us an unbiased estimate, while the method of counting the number of abnormals in a sample of fixed size n and dividing n by m would not do so. For, of course, the reciprocal of an unbiased estimate of a parameter is not an unbiased estimate of the reciprocal of the parameter.

Similarly

$$\begin{aligned}\overline{x^2} &= \sum_{n=m}^{\infty}\frac{(m-1)^2 w_n}{(n-1)^2}\\ &= (m-1)p^m q^{1-m}\sum_{n+m}^{\infty}\binom{n-2}{m-2}\frac{q^{n-1}}{n-1}\\ &= (m-1)p^m q^{1-m}\int_0^q\sum_{n=m}^{\infty}\binom{n-2}{m-2}t^{n-2}\,dt\\ &= (m-1)p^m q^{1-m}\int_0^q t^{m-2}(1-t)^{1-m}\,dt. \qquad (1)\end{aligned}$$

When m is small this can be integrated directly, e.g. for $m=4$,

$$\overline{x^2} = \frac{3p^2}{2q^2}\left(q-2p-\frac{2p^2\log p}{q}\right).$$

For larger values of m it is better to expand the integral in an infinite series. Put $u=\frac{pt}{q(1-t)}$, so that $t=\frac{qu}{p+qu}$, $dt=\frac{pq\,du}{(p+qu)^2}$.

Then

$$\begin{aligned}\overline{x^2} &= (m-1)p^2\int_0^1\frac{u^{m-2}\,du}{p+qu}\\ &= (m-1)p^2\int_0^1 u^{m-2}[1-q(1-u)]^{-1}\,du\\ &= (m-1)p^2\sum_{r=0}^{\infty}q^r\int_0^1 u^{m-2}(1-u)^r\,du\\ &= (m-1)p^2\sum_{r=0}^{\infty}\frac{(m-2)!\,r!\,q^r}{(m+r-1)!}\\ &= p^2\sum_{r=0}^{\infty}\frac{(m-1)!\,r!}{(m+r-1)!}q^r = p^2\sum_{r=0}^{\infty}\binom{m+r-1}{r}^{-1}q^r\\ &= p^2\left[1+\frac{q}{m}+\frac{2!\,q^2}{m(m+1)}+\frac{3!\,q^3}{m(m+1)(m+2)}+\dots\right].\end{aligned}$$

Hence the variance of x is

$$\sigma^2 = \overline{x^2} - p^2 = \frac{p^2 q}{m}\left[1 + \frac{2!\,q}{m+1} + \frac{3!\,q^2}{(m+1)(m+2)} + \dots\right]. \tag{2}$$

If in this equation we insert the estimated values of p and q, namely, $\frac{m-1}{n-1}$ and $\frac{n-m}{n-1}$, we find

$$\sigma^2 = \frac{m(n-m)}{n^2(n-1)}\left[1 + \frac{2(n-m)(n-3m)}{m^2 n^2} + O(m^{-3})\right]. \tag{3}$$

Thus $\sigma^2 = \frac{m(n-m)}{n^2(n-1)}$ is sufficiently accurate for all purposes and the classical value $\sigma^2 = \frac{m(n-m)}{n^3}$ sufficiently for most. These values can also be taken as the variances of p when m and n are given, though of course the exact values will depend on the prior probability distribution of p. It must be noted that if the variance is calculated from the estimated value of p it is approximately $\frac{p^2 q}{m-2}$, not $\frac{p^2 q}{m}$, so long as p is small. Since σ approximates to $p\sqrt{\frac{q}{m-2}}$, the standard error of p is a nearly constant fraction of the value of p when p is small, provided m is kept constant throughout a series of determinations. The higher moments can be obtained by a similar method, but as they involve multiple integrals, the series expansions are somewhat complicated.

Suppose now that the population sampled consists of several classes, and that a count is made until the number of the smallest class is m. Then the remaining $n-m$ consist of the other classes, and if their frequencies are q_1, q_2, q_3, etc., the sum being q, the expected numbers are $\frac{(n-m)q_1}{(n-1)q}$, etc. Hence the following rule may be laid down:

'If a sample is counted until the number of members of one class is m_1, those of the others being then m_2, m_3, etc., and the total n, then the estimate of the frequency of the first class is $\frac{m_1-1}{n-1}$, those of the others $\frac{m_2}{n-1}, \frac{m_3}{n-1}$, etc. The variances of these frequencies are approximately $m_r(n-m_r)\,n^{-3}$.'

A formal proof presents no difficulty.

Similar problems have of course been discussed, and there is a close analogy with problems concerning the duration of play when gambling (cf. Fieller, 1931). Here we have to find the probability that a player starting with a stake of £m, and with a probability p of winning £1 and a probability $1-p$ of losing it per game, will be ruined after n games. This could be generalized to cover the case where the probability of a win was r, of a draw q, and of a loss p; and the problem here considered would be the degenerate case where $r = 0$.

There is also an analogy with one of the methods in use for estimating the frequency of abnormal conditions in human families. These methods have been reviewed by Haldane (1938). If we find m abnormals in a family of n, the expected frequency (on certain assumptions) in a very large family produced by the same parents is not $\frac{m}{n}$ but $\frac{m-1}{n-1}$. For families of this type containing no abnormals would not be recorded, and therefore it is reasonable, as Weinberg (1927) pointed out, to take one of the abnormals as being merely a guarantee that the family is of a type including abnormals, and to base our frequency estimate on his or her sibs.

I have discussed this problem in some detail because I believe that if it is realized that

frequencies can be estimated just as accurately by counting up to a certain number of the rarest type as by counting a certain total, haematologists and others will be saved a good deal of needless effort. I have to thank Drs Kestelmann and Spurway for suggestions.

SUMMARY

If the frequency of an attribute is estimated by counting a sample until m members possessing this attribute are found, and if the total number in the sample is n, then $\frac{m-1}{n-1}$ is an unbiased estimate of the frequency, and its variance is very approximately $\frac{m(n-m)}{n^2(n-1)}$.

REFERENCES

FIELLER, E. C. (1931). The duration of play. *Biometrika*, **22**, 377.

GRÜNEBERG, H. (1942). The anaemia of flexed-tailed mice (*Mus musculus* L.). II. Siderocytes. *J. Genet.* **44**, 246.

HALDANE, J. B. S. (1938). The estimation of the frequencies of recessive conditions in man. *Ann. Eugen., Lond.*, **8**, 255.

WEINBERG, W. (1927). Mathematische Grundlagen der Probanden-methode. *Z. indukt. Abstamm- u. Vererb-Lehre*, **48**, 179.

THE MATHEMATICS OF A POPULATION COMPOSED OF k STATIONARY STRATA EACH RECRUITED FROM THE STRATUM BELOW AND SUPPORTED AT THE LOWEST LEVEL BY A UNIFORM ANNUAL NUMBER OF ENTRANTS

By H. L. SEAL

Consider a population subject to specified stochastic decremental forces operating at each attained age and supported by a uniform annual number of entrants distributed between ages α and β according to a known probability law. It is assumed that the population has existed for at least $\omega-\alpha$ years where ω represents the age, not necessarily finite, at which the expected number of individuals in the population becomes identically zero. The individuals of the, now stationary, population are subdivided into k strata by titular or other distinctions conferred without reference to characteristics influencing the incidence of the decremental forces. The numbers in each of the strata are determined by the action of a stochastic selective force operating on the individuals in stratum g ($g = 1, 2, \ldots, k-1$), with an intensity which depends only on the length of time the individual has spent in that stratum; if selected the individual moves up into stratum $(g+1)$. It is assumed that the total decremental and selective forces acting at each age and duration, respectively, are invariant with the passage of time.

Write μ_x for the total decremental force applicable stochastically throughout the population at exact age x; this function is supposed continuous and subject to $0 < \epsilon < \mu_x$. If ${}_tp_x$ equals the probability that an individual now aged x survives as a member of the population for at least t years then,

$${}_tp_x = \exp\left[-\int_0^t \mu_{x+\xi}\, d\xi\right] = \frac{\lambda_{x+t}}{\lambda_x} \quad \text{say} \quad (0 \leqslant t < \omega - x),$$

where $\lambda_\alpha > 0$ may be chosen arbitrarily.

Let $\lambda^{(g)}_{[\xi]}\, d\xi$, ($g = 1, 2, \ldots, k$) denote the expected number of entrants at any given moment into stratum g at exact age ξ; in particular, $\lambda^{(1)}_{[\xi]}$ is assumed to be a normalized function of bounded variation with $\lambda^{(1)}_{[\xi]} = 0$ $(0 \leqslant \xi < \alpha;\ \xi > \beta)$. Write $p^{(g)}_t$ for the probability that an individual, if alive, remains unselected for at least t years after entry into stratum g, ($g = 1, 2, \ldots, k$; $p^{(k)}_t = 1$); $p^{(g)}_t$ is thus bounded and monotonically decreasing, and at points of discontinuity it is to be defined by

$$p^{(g)}_t = \tfrac{1}{2}\{p^{(g)}_{t-0} + p^{(g)}_{t+0}\}.$$

Writing $q^{(g)}_t = 1 - p^{(g)}_t$, $q^{(g)}_t$ is a normalized function of bounded variation in $(0, \omega)$.

It is intended to derive expressions in terms of λ_x, $\lambda^{(1)}_{[x]}$ and $p^{(g)}_t$ ($g = 1, 2, \ldots, k$), supposed known, for:

(i) $\lambda^{(g)}_{[x]}\, dx$ ($g = 1, 2, \ldots, k$), the expected number of entrants at any given moment into stratum g at exact age x,

(ii) $\Lambda^{(g)}$ ($g = 1, 2, \ldots, k$), the expected total number of individuals in stratum g at any moment of time.

The expected number of entrants at age x into grade $(g+1)$ is defined to be the aggregate of all selectees for this grade deriving from entrants into grade g at all ages below x; hence, formally,

$$\lambda_{[x]}^{(g+1)} = \int_0^x \lambda_{[x-\xi]}^{(g)}\, {}_\xi p_{x-\xi}\, dq_\xi^{(g)} \quad (g = 1, 2, \ldots, k-1),$$

i.e.

$$\frac{\lambda_{[x]}^{(g+1)}}{\lambda_x} = \int_0^x \frac{\lambda_{[x-\xi]}^{(g)}}{\lambda_{x-\xi}}\, dq_\xi^{(g)}. \tag{1}$$

The integral on the right-hand side of (1) exists when $g = 1$ and, after defining $\lambda_{[x]}^{(2)}$ suitably at points of discontinuity, $\lambda_{[x]}^{(2)}$ is a normalized function of bounded variation in $(0, \omega)$ (Widder, 1941, Ch. II). By induction $\lambda_{[x]}^{(g)}$ $(g = 3, 4, \ldots, k)$, are normalized functions of bounded variation in $(0, \omega)$.

The expected total of individuals in strata $g+1, g+2, \ldots, k$, is composed of all the survivors from entrants at various ages into grade $(g+1)$; thus

$$\begin{aligned}\sum_{r=g+1}^{k} \Lambda^{(r)} &= \int_0^\omega \lambda_{[x]}^{(g+1)}\, dx \int_0^{\omega-x} {}_t p_x\, dt \\ &= \int_0^\omega \frac{\lambda_{[x]}^{(g+1)}}{\lambda_x}\, dx \int_x^\omega \lambda_\xi\, d\xi \\ &= \int_0^\omega \lambda_\xi\, d\xi \int_0^\xi \frac{\lambda_{[x]}^{(g+1)}}{\lambda_x}\, dx \quad \text{by Dirichlet's integral theorem,}\end{aligned}$$

or

$$\sum_{r=g+1}^{k} \Lambda^{(r)} = \int_0^\omega \lambda_\xi\, d\xi \int_0^\xi (\xi - x)\, d\frac{\lambda_{[x]}^{(g+1)}}{\lambda_x}. \tag{2}$$

Now write, $s = \sigma + i\tau$,

$$f_r(s) = \begin{cases} \displaystyle\int_0^\infty e^{-st}\, dq_t^{(r)} & (r = 1, 2, \ldots, k), \\ \displaystyle\int_0^\infty e^{-st}\, d\frac{\lambda_{[t]}^{(1)}}{\lambda_t} & (r = 0), \end{cases} \tag{3}$$

then (Widder, 1941) if the integrals defining $f_r(s)$ all converge for $\sigma = \sigma_0 > 0$, the rule of multiplication of Laplace transforms results in

$$\int_0^\infty e^{-st}\, d\frac{\lambda_{[t]}^{(g+1)}}{\lambda_t} = \prod_{r=0}^{g} f_r(s) = \phi_g(s) \quad \text{say} \quad (g = 0, 1, 2, \ldots, k-1)$$

and, inverting, $(c > \sigma_0)$,

$$\frac{\lambda_{[x]}^{(g+1)}}{\lambda_x} = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{e^{sx}}{s} \phi_g(s)\, ds. \tag{4}$$

In a similar manner, if

$$Q_\xi^{(g+1)} = \int_0^\xi (\xi - x)\, d\frac{\lambda_{[x]}^{(g+1)}}{\lambda_x} \quad (g = 0, 1, 2, \ldots, k-1)$$

$$\int_0^\infty e^{-st}\, dQ_t^{(g+1)} = \prod_{r=0}^{g} f_r(s) \int_0^\infty e^{-st}\, dt = \frac{1}{s} \phi_g(s)$$

and $$Q_\xi^{(g+1)} = \lim_{T\to\infty} \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{e^{s\xi}}{s^2}\phi_g(s)\,ds \quad (c>\sigma_0),$$

so that $$\sum_{r=g+1}^{k} \Lambda^{(r)} = \frac{1}{2\pi i}\int_0^\omega \lambda_\xi\,d\xi \int_{c-i\infty}^{c+i\infty} \frac{e^{s\xi}}{s^2}\phi_g(s)\,ds. \tag{5}$$

Example (i). Let $(r = 1, 2, \ldots, k-1)$,

$$-\frac{1}{p_t^{(r)}}\frac{dp_t^{(r)}}{dt} = \begin{cases} c_r & a_r \leqslant t \leqslant b_r, \\ 0 & t < a_r, t > b_r, \end{cases}$$

then $$\frac{dq_t^{(r)}}{dt} = -\frac{dp_t^{(r)}}{dt} = \begin{cases} c_r e^{-c_r(t-a_r)} & a_r \leqslant t \leqslant b_r, \\ 0 & t < a_r,\ t > b_r, \end{cases}$$

and thus $(r = 2, 3, \ldots, k-1)$,

$$f_r(s) = \int_0^\infty e^{-st}\,dq_t^{(r)} = c_r e^{c_r a_r}\int_{a_r}^{b_r} e^{-(s+c_r)t}\,dt = \frac{c_r}{s+c_r}(e^{-sa_r} - R_r e^{-sb_r}),$$

where $R_r = e^{-c_r(b_r-a_r)}$.

It is further assumed that

$$\frac{\lambda_{[t]}^{(1)}}{\lambda_t} = \begin{cases} 1 & t = \alpha, \\ 0 & t \neq \alpha, \end{cases}$$

and since in this case relation (1) does not hold for $g = 1$, $f_0(s)$ must be replaced by unity and $f_1(s)$ given by

$$f_1(s) = \int_0^\infty e^{-st}\,d\frac{\lambda_{[t]}^{(2)}}{\lambda_t},$$

where $$\lambda_{[x]}^{(2)} = \lambda_{[\alpha]}^{(1)}{}_{x-\alpha}p_\alpha \frac{dq_{x-\alpha}^{(1)}}{dx} \quad (\alpha \leqslant x < \omega),$$

i.e. $$\frac{\lambda_{[x]}^{(2)}}{\lambda_x} = \frac{dq_{x-\alpha}^{(1)}}{dx},$$

so that

$$f_1(s) = \int_0^\infty e^{-st}\,d\frac{\lambda_{[t]}^{(2)}}{\lambda_t} = s\int_0^\infty e^{-st}\,dq_{t-\alpha}^{(1)} = se^{-s\alpha}\int_0^\infty e^{-st}\,dq_t^{(1)} = \frac{se^{-s\alpha}c_1}{s+c_1}(e^{-sa_1} - R_1 e^{-sb_1}).$$

Thus $$\phi_g(s) = \frac{C_g s e^{-s\alpha}}{\prod_{r=1}^{g}(s+c_r)}\prod_{r=1}^{g}(e^{-sa_r} - R_r e^{-sb_r}), \quad \text{where } C_g = \prod_{r=1}^{g} c_r$$

$$= \frac{C_g s e^{-s\alpha}}{\prod_{r=1}^{g}(s+c_r)}\sum_{\nu=1}^{2^g} T_\nu e^{-sW_g(\nu)},$$

where $W_g(\nu)$ assumes in succession all the possible sum-combinations of g of the quantities a_r, b_r $(r = 1, 2, \ldots, g)$, without using an a and a b with the same suffix. The coefficient T_ν is equal to the product of those $-R_r$'s corresponding to the b_r's used in the formation of $W_g(\nu)$.

Hence

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{s\xi}}{s^2}\phi_g(s)\,ds = \frac{C_g}{2\pi}\sum_{\nu=1}^{2^g} T_\nu e^{(\xi-\alpha-W_g(\nu))c}\int_{-\infty}^{\infty}\frac{e^{it(\xi-\alpha-W_g(\nu))}}{(c+it)\prod\limits_{r=1}^{g}(c_r+c+it)}\,dt$$

$$= \frac{C_g}{2\pi}\sum_{\nu=1}^{2^g} T_\nu e^{(\xi-\alpha-W_g(\nu))c}\int_{-\infty}^{\infty} e^{it(\xi-\alpha-W_g(\nu))}\sum_{r=0}^{g}\frac{Q_r}{c_r+c+it}\,dt \quad (c_0=0,\ c_r\neq c_j;\ r,j=1,2,\ldots,g)$$

$$= C_g\sum_{\nu=1}^{2^g} T_\nu\sum_{r=0}^{g} Q_r e^{-c_r(\xi-\alpha-W_g(\nu))} \quad (\xi > \alpha + W_g(\nu)),$$

where

$$Q_r = \Big\{\prod_{j\neq r}(c_j-c_r)\Big\}^{-1} \quad (r = 0,1,2,\ldots,g).$$

(A similar type of expression may be found for $\lambda_{[x]}^{(g+1)}/\lambda_x$.)

And thus, by (5),

$$\sum_{r=g+1}^{k} A^{(r)} = C_g\int_0^{\omega}\lambda_x\left\{\sum_{\nu=1}^{2^g}T_\nu\sum_{r=0}^{g}Q_r e^{-c_r(x-\alpha-W_g(\nu))}\right\}dx$$

$$= C_g\sum_{\nu=1}^{2^g}T_\nu\sum_{r=0}^{g}Q_r e^{c_r(\alpha+W_g(\nu))}\,\overline{N}^{(c_r)}_{\alpha+W_g(\nu)},$$

where

$$\overline{N}^{(c_r)}_{\alpha+W_g(\nu)} = \int_0^{\omega-\alpha-W_g(\nu)} e^{-c_r(\alpha+W_g(\nu)+t)}\lambda_{\alpha+W_g(\nu)+t}\,dt.$$

The above relation holds provided $c_r \neq c_j$ $(r, j = 1, 2, \ldots, g)$; if, on the other hand, $c_1 = c_2 = c_3 = \ldots = c_g$, it may be shown that

$$\sum_{r=g+1}^{k} A^{(r)} = \sum_{\nu=1}^{2^g} T_\nu\left\{\overline{N}^{(0)}_{\alpha+W_g(\nu)} - \sum_{r=1}^{g}\frac{c_1^{r-1}}{(r-1)!}\int_0^{\omega-\alpha-W_g(\nu)} t^{r-1}e^{-c_1 t}\lambda_{\alpha+W_g(\nu)+t}\,dt\right\}.$$

The case where some of the c's are equal presents no particular difficulties.

The results obtained have a close analogy with those appropriate to a problem in the theory of radioactive transformations (Bateman, 1910).

Example (ii). Let

$$\frac{\lambda_{[t]}^{(1)}}{\lambda_t} = \begin{cases} 1 & \alpha < t < \beta, \\ \tfrac{1}{2} & t = \alpha,\ t = \beta, \\ 0 & 0 \leqslant t < \alpha,\ t > \beta, \end{cases}$$

and $(r = 1, 2, \ldots, k-1)$

$$-\frac{dp_t^{(r)}}{dt} = \frac{dq_t^{(r)}}{dt} = \begin{cases} \dfrac{1}{\Gamma(b_r)}e^{-a(t-c_r)}(t-c_r)^{b_r-1} & t > c_r \quad (a, b_r > 0), \\ 0 & t < c_r \quad (\alpha + c_1 > \beta). \end{cases}$$

Then

$$f_0(s) = e^{-s\alpha} - e^{-s\beta}$$

and

$$f_r(s) = \frac{e^{-sc_r}}{(s+a)^{b_r}} \quad (r = 1, 2, \ldots, k-1),$$

and thus $(g = 0, 1, 2, \ldots, k-1)$

$$\phi_g(s) = \prod_{r=0}^{g} f_r(s) = \frac{e^{-sC_g^0} - e^{-sC_g^1}}{(s+a)^{B_g}},$$

where

$$C_g^0 = \alpha + \sum_{r=1}^{g} c_r, \quad B_g = \sum_{r=1}^{g} b_r \quad \text{and} \quad C_g^1 = C_g^0 - \alpha + \beta \quad (\text{Obs. } C_0^0 = \alpha,\ C_0^1 = \beta,\ B_0 = 0).$$

Hence

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{s\xi}}{s^2}\phi_g(s)\,ds=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{e^{s(\xi-C_g^0)}-e^{s(\xi-C_g^1)}}{s^2(s+a)^{B_g}}\,ds$$

$$=\frac{e^{(\xi-C_g^0)c}}{2\pi}\int_{-\infty}^{\infty}\frac{e^{it(\xi-C_g^0)}}{(c+it)^2\,(a+c+it)^{B_g}}\,dt-\frac{e^{(\xi-C_g^1)c}}{2\pi}\int_{-\infty}^{\infty}\frac{e^{it(\xi-C_g^1)}}{(c+it)^2\,(a+c+it)^{B_g}}\,dt.$$

But
$$\frac{e^{xc}}{2\pi}\int_{-\infty}^{\infty}\frac{e^{itx}}{(c+it)^2\,(a+c+it)^{B_g}}\,dt=\begin{cases}\dfrac{1}{\Gamma(B_g)}\displaystyle\int_0^x e^{-ay}\,y^{B_g-1}(x-y)\,dy & x>0,\\ 0 & x<0,\end{cases}$$

(see, e.g. § 13·8 of Bochner, 1932) and

$$\int_{C_g^0}^{\omega}\lambda_x dx\int_0^{x-C_g^0}e^{-ay}\,y^{B_g-1}(x-C_g^0-y)\,dy=\int_0^{\omega-C_g^0}\lambda_{\xi+C_g^0}d\xi\int_0^{\xi}e^{-ay}\,y^{B_g-1}(\xi-y)\,dy$$

$$=\int_0^{\omega-C_g^0}e^{-ay}\,y^{B_g-1}\,dy\int_y^{\omega-C_g^0}(\xi-y)\,\lambda_{\xi+C_g^0}d\xi.$$

Writing
$$\bar{S}_z^0=\int_0^{\omega-z}t\lambda_{z+t}dt$$

(cp. Steffensen, 1934), there results from (5)

$$\sum_{r=g+1}^{k}\Lambda^{(r)}=\frac{1}{\Gamma(B_g)}\int_0^{\omega-C_g^0}e^{-ay}\,y^{B_g-1}\bar{S}_{y+C_g^0}^0\,dy-\frac{1}{\Gamma(B_g)}\int_0^{\omega-C_g^1}e^{-ay}\,y^{B_g-1}\bar{S}_{y+C_g^1}^0\,dy.$$

I am very grateful to Dr Stefan Vajda for a number of useful suggestions in connexion with the preceding analysis.

REFERENCES

BATEMAN, H. (1910). The solution of a system of differential equations occurring in the theory of radio-active transformations. *Proc. Camb. Phil. Soc.* **15**, 423.

BOCHNER, S. (1932). *Vorlesungen über Fouriersche Integrale.* Leipzig.

STEFFENSEN, J. F. (1934). *Forsikringsmatematik.* Copenhagen.

WIDDER, D. V. (1941). *The Laplace Transform.* Princeton.

MOMENTS OF r AND χ^2 FOR A FOURFOLD TABLE IN THE ABSENCE OF ASSOCIATION

By J. B. S. HALDANE, F.R.S.

Fourfold contingency tables are in constant use, and it is of interest to calculate the moments of χ^2 derived from them when one or more of the expectations are small, even though the exact method of Fisher (1936) and the table of Fisher & Yates (1938) render this less important than would once have been the case. Further, Kendall's (1942) use of the product-moment correlation of a fourfold table in connexion with rank correlation has given it a new interest. As the third moment of χ^2 and the sixth moment of r are readily derivable as special cases from the formulae of Haldane (1945), it seems worth while to give them, along with the third and fifth moments of r.

Consider the table

a	b	L
c	d	l
M	m	S

$$\chi^2 = \frac{(ad-bc)^2 S}{LlMm}, \quad \text{and} \quad r = \frac{ad-bc}{(LlMm)^{\frac{1}{2}}} = S^{-\frac{1}{2}}\chi.$$

If L and l or M and m are samples of the same population falling into two classes, then $\bar{r} = 0$, and $\overline{\chi^2} = 1$ approximately. The exact values of the first three moments of χ^2 are readily derived from equations (1) of Haldane (1945), putting

$$n = 2, \quad k = \frac{S^2}{Ll}, \quad R_1 = \frac{S}{Mm}, \quad R_2 = \frac{S^2}{M^2m^2} - \frac{2}{Mm}.$$

Let $LlMm = \lambda^2$, $Ll + Mm = \mu$, $(L-l)(M-m) = \nu$. Then

$$\overline{\chi^2} = \frac{S}{S-1},$$

$$\overline{\chi^4} = \frac{S^2[3(S+6)\lambda^2 - 6S^2\mu + S^3(S+1)]}{\lambda^2(S-1)(S-2)(S-3)},$$

$$\overline{\chi^6} = \frac{S^3}{\lambda^4(S-1)(S-2)(S-3)(S-4)(S-5)}[5(3S^2+86S+120)\lambda^4 - 10S^2(13S+60)\lambda^2\mu$$
$$+ 120S^4\mu^2 - 5S^2(5S^2+87S+60)\lambda^2 - 30S^5(S+3)\mu + S^5(S+1)(S^2+15S-4)],$$

$$\kappa_1 = \mu_1' = \frac{S}{S-1},$$

$$\kappa_2 = \mu_2 = \frac{S^2}{\lambda^2(S-1)(S-2)(S-3)}[2(S-1)^{-1}(S^2+10S-12)\lambda^2 - 6S^2\mu + S^3(S+1)],$$

$$\kappa_3 = \mu_3 = \frac{S^3}{\lambda^4(S-1)(S-2)(S-3)(S-4)(S-5)}[8(S-1)^{-2}(S^4+51S^3+22S^2-308S+240)\lambda^4$$
$$- 8(S-1)^{-1}S^2(14S^2+79S-120)\lambda^2\mu + 120S^4\mu^2$$
$$- 2(S-1)^{-1}S^3(14S^3+193S^2-51S-120)\lambda^2 - 30S^5(S+3)\mu$$
$$+ S^5(S+1)(S^2+15S-4)]. \qquad (1)$$

The odd moments of r can be calculated as follows. To obtain the mean value of an odd power of $(ad-bc)$ we use the operator Δ. This permutes the indices of a and d, and also those of b and c, without change of sign. It also permutes the indices (normal and factorial) of a and d with those of b and c, the sign being changed. No term is repeated if obtained more than once, and all are added together. Thus:

$$\Delta(a^3d^2) = a^3d^2+a^2d^3-b^3c^2-b^2c^3,$$

$$\Delta(a^3bd^2) = a^3bd^2+a^2bd^3+a^3cd^2+a^2cd^3-ab^3c^2-b^3c^2d-ab^2c^3-b^2c^3d.$$

Now the mean value of a product of factorials of a, b, c, d is

$$\begin{aligned} E[a^{(\alpha)}b^{(\beta)}c^{(\gamma)}d^{(\delta)}] &= E[a^{(\alpha+\epsilon)}b^{(\beta-\epsilon)}c^{(\gamma-\epsilon)}d^{(\delta-\epsilon)}] \\ &= L^{(\alpha+\beta)}l^{(\gamma+\delta)}M^{(\alpha+\gamma)}m^{(\beta+\delta)}S^{-(\alpha+\beta+\gamma+\delta)}, \end{aligned}$$

where $a^{(\alpha)} \equiv a(a-1)(a-2)\dots(a-\alpha+1)$, and so on (Haldane, 1940). It follows that

$$E\Delta[a^{(\alpha)}b^{(\beta)}c^{(\gamma)}d^{(\delta)}] = 0, \quad \text{if} \quad \alpha+\beta=\gamma+\delta \text{ or } \alpha+\gamma=\beta+\delta,$$

and

$$\begin{aligned} E\Delta[a^{(\alpha)}d^{(\delta)}] &= S^{-(\alpha+\delta)}[L^{(\alpha)}l^{(\delta)}-L^{(\delta)}l^{(\alpha)}][M^{(\alpha)}m^{(\delta)}-M^{(\delta)}m^{(\alpha)}], \\ E\Delta[a^{(\alpha)}bd^{(\delta)}] &= S^{-(\alpha+\delta+1)}[\{L^{(\alpha+1)}l^{(\delta)}-L^{(\delta)}l^{(\alpha+1)}\}\{M^{(\alpha)}m^{(\delta+1)}-M^{(\delta+1)}m^{(\alpha)}\} \\ &\qquad +\{L^{(\alpha)}l^{(\delta+1)}-L^{(\delta+1)}l^{(\alpha)}\}\{M^{(\alpha+1)}m^{(\delta)}-M^{(\delta)}m^{\alpha+1}\}], \\ E\Delta[a^{(\alpha)}bd^{(\alpha)}] &= -E\Delta[a^{(\alpha+1)}d^{(\alpha)}]. \end{aligned}$$

Thus

$$\begin{aligned} (ad-bc)^3 &= \Delta(a^3d^3-3a^2bcd^2) \\ &= \Delta[a^{(3)}d^{(3)}-3a^{(2)}bcd^{(2)}+3a^{(3)}d^{(2)}-3a^{(2)}bcd+a^{(3)}d+9a^{(2)}d^{(2)}+3a^{(2)}d+ad]. \end{aligned}$$

So

$$\begin{aligned} E[(ad-bc)^3] &= E\Delta[a^{(3)}d+3a^{(2)}d], \quad \text{the other terms vanishing,} \\ &= S^{-(4)}[L^{(3)}l-Ll^{(3)}][M^{(3)}m-Mm^{(3)}]+3S^{-(3)}[L^{(2)}l-Ll^{(2)}][M^{(2)}m-Mm^{(2)}] \\ &= (S-1)^{-(2)}\lambda^2\nu. \end{aligned}$$

In general, a^α must be expanded in factorials as a polynomial whose coefficients are the initial differences of the powers of integers, divided by the appropriate factorials.

$$\begin{aligned} (ad-bc)^5 &= \Delta(a^5d^5-5a^4bcd^4+10a^3b^2c^2d^3) \\ &= \Delta[a^{(5)}d^{(5)}-5a^{(4)}bcd^{(4)}+10a^{(3)}b^{(2)}c^{(2)}d^{(3)} \\ &\quad +10\{a^{(5)}d^{(4)}-3a^{(4)}bcd^{(3)}+3a^{(3)}b^{(2)}c^{(2)}d^{(2)}+a^{(3)}b^{(2)}cd^{(3)}\} \\ &\quad +5\{5a^{(5)}d^{(3)}+20a^{(4)}d^{(4)}-7a^{(4)}bcd^{(2)}-34a^{(3)}bcd^{(3)}+2a^{(3)}b^{(2)}c^{(2)}d+6a^{(3)}b^{(2)}cd^{(2)} \\ &\quad +18a^{(2)}b^{(2)}c^{(2)}d^{(2)}\} \\ &\quad +5\{3a^{(5)}d^{(2)}+50a^{(4)}d^{(3)}-a^{(4)}bcd^{(2)}-36a^{(3)}bcd^{(2)}-12a^{(2)}b^{(2)}c^{(2)}d+2a^{(3)}b^{(2)}cd\} \\ &\quad +\{a^{(5)}d+150a^{(4)}d^{(2)}+625a^{(3)}d^{(3)}-20a^{(3)}bcd-165a^{(2)}bcd^{(2)}+30a^{(2)}b^{(2)}cd\} \\ &\quad +5\{2a^{(4)}d+75a^{(3)}d^{(2)}-3a^{(2)}bcd\}+25\{a^{(3)}d+9a^{(2)}d^{(2)}\}+15a^{(2)}d]. \end{aligned}$$

So

$$\begin{aligned} E[(ad-bc)^5] &= E\Delta[10\{a^{(5)}d^{(2)}+a^{(4)}d^{(3)}+a^{(4)}bd^{(2)}\}+\{a^{(5)}d+130a^{(4)}d^{(2)}\} \\ &\qquad +10\{a^{(4)}d+36a^{(3)}d^{(2)}\}+25a^{(3)}d+15a^{(2)}d] \\ &= 10S^{-(7)}[\{L^{(5)}l^{(2)}-L^{(2)}l^{(5)}\}\{M^{(5)}m^{(2)}-M^{(2)}m^{(5)}\}+\{L^{(4)}l^{(3)}-L^{(3)}l^{(4)}\}\{M^{(4)}m^{(3)}-M^{(3)}m^{(4)}\} \\ &\quad +\{L^{(5)}l^{(2)}-L^{(2)}l^{(5)}\}\{M^{(4)}m^{(3)}-M^{(3)}m^{4}\}+\{L^{(4)}l^{(3)}-L^{(3)}l^{(4)}\}\{M^{(5)}m^{(2)}-M^{(2)}m^{(5)}\}] \\ &\quad +S^{-(6)}[\{L^{(5)}l-Ll^{(5)}\}\{M^{(5)}m-Mm^{(5)}\}+130\{L^{(4)}l^{(2)}-L^{(2)}l^{(4)}\}\{M^{(4)}m^{(2)}-M^{(2)}m^{(4)}\}] \\ &\quad +10S^{-(5)}[\{L^{(4)}l-Ll^{(4)}\}\{M^{(4)}m-Mm^{(4)}\}+36\{L^{(3)}l^{(2)}-L^{(2)}l^{(3)}\}\{M^{(3)}m^{(2)}-M^{(2)}m^{(3)}\}] \\ &\quad +25S^{-(4)}[L^{(3)}l-Ll^{(3)}][M^{(3)}m-Mm^{(3)}]+15[L^{(2)}l-Ll^{(2)}][M^{(2)}m-Mm^{(2)}] \end{aligned}$$

$$
\begin{aligned}
&= 10S^{-(7)}\lambda^2\nu[\lambda^2-(S-1)\mu+(S-1)^2](S-5)^2(S-6)^2\\
&\quad + S^{-(6)}\lambda^2\nu[134\lambda^2-2(S^2+60S-55)\mu+S^4-10S^3+175S^2-360S+230](S-5)^2\\
&\quad + 10S^{-(5)}\lambda^2\nu[37\lambda^2-(S^2+30S-25)\mu+S^4-12S^3+94S^2-204S+157]\\
&\quad + 25S^{-(4)}\lambda^2\nu(S-3)^2+15S^{-(3)}\lambda^2\nu\\
&= (S-1)^{-(4)}\lambda^2\nu[2(5S+12)\lambda^2-12S^2\mu+S^3(S+5)].
\end{aligned}
$$

Hence the moments of r are

$$
\begin{aligned}
\mu_1' &= 0,\\
\mu_2 &= \frac{1}{S-1},\\
\mu_3 &= \frac{\nu}{(S-1)(S-2)},\\
\mu_4 &= \frac{3(S+6)\lambda^2-6S^2\mu+S^3(S+1)}{\lambda^2(S-1)(S-2)(S-3)},\\
\mu_5 &= \frac{\nu[2(5S+12)\lambda^2-12S^2\mu+S^3(S+5)]}{\lambda^3(S-1)(S-2)(S-3)(S-4)},\\
\mu_6 &= [5(3S^2+86S+120)\lambda^4-10S^2(13S+60)\lambda^2\mu+120S^4\mu^2-5S^2(5S^2+87S+60)\lambda^2\\
&\quad -30S^5(S+3)\mu+S^5(S+1)(S^2+15S-4)]\div\lambda^4(S-1)(S-2)(S-3)(S-4)(S-5),\\
\kappa_4 &= \frac{\dfrac{6(5S-6)\lambda^2}{S-1}-6S^2\mu+S^3(S+1)}{\lambda^2(S-1)(S-2)(S-3)(S-4)},\\
\kappa_5 &= \frac{\nu\left[\dfrac{12(7S-12)\lambda^2}{S-1}-12S^2\mu+S^3(S+5)\right]}{\lambda^3(S-1)(S-2)(S-3)(S-4)}.
\end{aligned}
$$

The odd moments vanish if $L = l$, or $M = m$, i.e. if the samples are equal, or the class frequencies 50%. κ_4 is negative if L and l, or M and m, are nearly equal, but becomes positive if both a class frequency and a sample are small.

These expressions may be used for accurate tests of the significance of observed values of χ^2 or r, or, which is more difficult by the methods mentioned above, the significance of a series of values of these parameters.

REFERENCES

FISHER, R. A. (1936). The Logic of inductive inference. *J. Roy. Statist. Soc.* **98**, 39–82.

FISHER, R. A. & YATES, F. (1938). *Statistical Tables for Biological, Medical and Agricultural Research.* Edinburgh.

HALDANE, J. B. S. (1937). The exact value of the moments of the distribution of χ^2, used as a test of goodness of fit when expectations are small. *Biometrika*, **29**, 389–91.

HALDANE, J. B. S. (1940). The mean and variance of χ^2, when used as a test of homogeneity, when expectations are small. *Biometrika*, **31**, 346–55.

HALDANE, J. B. S. (1945). The use of χ^2 as a test of homogeneity in a $(n \times 2)$-fold table when expectations are small. *Biometrika*, **33**, 234.

KENDALL, M. G. (1942). Partial rank correlation. *Biometrika*, **32**, 277–83.

THE USE OF χ^2 AS A TEST OF HOMOGENEITY IN A $(n \times 2)$-FOLD TABLE WHEN EXPECTATIONS ARE SMALL

BY J. B. S. HALDANE, F.R.S.

The χ^2 test has proved so useful that any extension of its field of application is likely to be of value. Its use is at present deprecated when the expectation in any group is small. This may occur either because samples are small, or because one of the classes expected in each sample is rare. Fisher (1941) recommends 5 as a lower limit for expectations. It will be shown that the test may still be applied even when the expectation is less than unity. Where the number of classes is large, we may still sometimes use the ordinary tables. Otherwise the value of P must be calculated in each particular case. Haldane (1937) showed how this could be done when χ^2 is used as a test of goodness of fit. But, particularly in genetical work, it has found its greatest use as a test of homogeneity. The formulae required for a $(m \times n)$-fold table are very cumbrous. Those for a $(n \times 2)$-fold table are developed below. The value of the variance has already been given (Haldane, 1940). That of the third moment is now calculated.

Consider a set of n samples of $s_1, s_2, \ldots, s_i, \ldots, s_n$ individuals. In the ith sample let there be a_i individuals of class X and b_i of class Y. Let $\Sigma a_i = A$, $\Sigma b_i = B$, $\Sigma s_i = S$, so that the table is

a_1 b_1	a_2 b_2	$\ldots$ $\ldots$	a_i b_i	$\ldots$ $\ldots$	a_n b_n	A B
s_1	s_2	$\ldots$	s_i	$\ldots$	s_n	S

Let $$\Sigma s_i^{-1} = R_1, \quad \Sigma s_i^{-2} = R_2, \quad S^2/AB = k.$$

Then it can readily be shown that $\chi^2 = S\left(1 - \dfrac{Sx}{AB}\right)$, where $x = \Sigma\left(\dfrac{a_i b_i}{s_i}\right)$. It also follows at once from the argument of Haldane (1940, p. 347) that in a homogeneous population the expectation of $a_i^{(\alpha)} b_i^{(\beta)}$ is $S^{-(\alpha+\beta)} A^{(\alpha)} B^{(\beta)} s_i^{(\alpha+\beta)}$, where $y^{(\alpha)} = y(y-1)(y-2)\ldots(y-\alpha+1)$. Hence

$$\bar{x} = S^{-(2)} AB(S-n),$$

$$\overline{x^2} = S^{-(4)} A^2 B^2 [S^2 - 2(n+2)S + n(n+10) - 6R_1 + kS^{-2} \times \{-(n^2+2n-2)S + n(n-2) + R_1 S(S+1)\}].$$

The proof is similar to that which follows for $\overline{x^3}$, in which multiple summations are taken over all unequal sets of values of i, j, and i, j, k.

$$\begin{aligned} x^3 &= \Sigma s_i^{-3} a_i^3 b_i^3 + 3\Sigma\Sigma s_i^{-2} a_i^2 b_i^2 s_j^{-1} a_j b_j + 6\Sigma\Sigma\Sigma s_i^{-1} a_i b_i s_j^{-1} a_j b_j s_k^{-1} a_k b_k \\ &= \Sigma s_i^{-3}[a_i^{(3)} b_i^{(3)} + 3a_i^{(3)} b_i^{(2)} + 3a_i^{(2)} b_i^{(3)} + a_i^{(3)} b_i + 9a_i^{(2)} b_i^{(2)} + a_i b_i^{(3)} + 3a_i^{(2)} b_i + 3a_i b_i^{(2)} + a_i b_i] \\ &\quad + 3\Sigma\Sigma s_i^{-2} s_j^{-1}[a_i^{(2)} b_i^{(2)} a_j b_j + a_i^{(2)} b_i a_j b_j + a_i b_i^{(2)} a_j b_j + a_i b_i a_j b_j] \\ &\quad + 6\Sigma\Sigma\Sigma s_i^{-1} a_i b_i s_j^{-1} a_j b_j s_k^{-1} a_k b_k. \end{aligned}$$

Hence $\overline{x^3} = S^{-(6)}A^{(3)}B^{(3)}[\Sigma s_i^{-3}s_i^{(6)} + 3\,\Sigma\Sigma s_i^{-2}s_i^{(4)}s_j^{-1}s_j^{(2)} + 6\Sigma\Sigma\Sigma s_i^{-1}s_i^{(2)}s_j^{-1}s_j^{(2)}s_k^{-1}s_k^{(2)}]$

$$+3S^{-(5)}[A^{(3)}B^{(2)}+A^{(2)}B^{(3)}]\,[\Sigma s_i^{-3}s_i^{(5)} + \Sigma\Sigma s_i^{-2}s_i^{(3)}s_j^{-1}s_j^{(2)}]$$

$$+S^{-(4)}[A^{(3)}B + 9A^{(2)}B^{(2)} + AB^{(3)}]\,\Sigma s_i^{-3}s_i^{(4)} + 3S^{-(4)}A^{(2)}B^{(2)}\,\Sigma\Sigma s_i^{-2}s_i^{(2)}s_j^{-2}s_j^{(2)}$$

$$+3S^{-(3)}[A^{(2)}B+AB^{(2)}]\,\Sigma s_i^{-3}s_i^{(3)} + s_i^{-(2)}AB\,\Sigma s_i^{-3}s_i^{(2)}$$

$$= S^{-(6)}AB(AB-S+1)\,(AB-2S+4)\,[\Sigma(s_i^3-15s_i^2+85s_i-225+274s_i^{-1}-120s_i^{-2})$$

$$+3\Sigma\Sigma(s_i^2-6s_i+11-6s_i^{-1})\,(s_j-1)+6\Sigma\Sigma\Sigma(s_i-1)\,(s_j-1)\,(s_k-1)]$$

$$+3S^{-(5)}AB(AB-S+1)\,(S-4)\,[\Sigma(s_i^2-10s_i+35-50s_i^{-1}$$

$$+24s_i^{-2})+\Sigma\Sigma(s_i-3+2s_i^{-1})\,(s_j-1)]$$

$$+S^{-(4)}AB(7AB+S^2-12S+13)\,\Sigma(s_i-6+11s_i^{-1}-6s_i^{-2})$$

$$+3S^{-(4)}AB(AB-S+1)\,\Sigma\Sigma(1-s_i^{-1})\,(s_j-1)$$

$$+3S^{-(3)}AB(S-2)\,\Sigma(1-3s_i^{-1}+2s_i^{-2})+S^{-(2)}AB\Sigma(s_i^{-1}-s_i^{-2}).$$

But $\Sigma(s_i^3-3s_i^2+3s_i-1)+3\Sigma\Sigma(s_i^2-2s_i+1)\,(s_j-1)+6\Sigma\Sigma\Sigma(s_i-1)\,(s_j-1)\,(s_k-1)$

$$=[\Sigma(s_i-1)]^3=(S-n)^3,$$

$$\Sigma(s_i^2-2s_i+1)+2\Sigma\Sigma(s_i-1)\,(s_j-1)=[\Sigma(s_i-1)]^2=(S-n)^2,$$

$$\Sigma(s_i-2+s_i^{-1})+2\Sigma\Sigma(1-s_i^{-1})\,(s_j-1)=\Sigma(s_i-1)\,\Sigma(1-s_i^{-1})=(S-n)\,(n-R_1).$$

So $\overline{x^3} = S^{-(6)}AB[A^2B^2-(3S-5)\,AB+2(S-1)\,(S-2)]$

$$\times[(S-n)^3-12(S-n)^2+18(S-n)\,(n-R_1)+40S-176n+256R_1-120R_2]$$

$$+3S^{-(4)}AB(AB-S+1)\,[(S-n)^2-(S-n)\,(n-R_1)-7S+32n-49R_1+24R_2]$$

$$+S^{-(4)}AB(7AB+S^2-12S+13)\,(S-6n+11R_1-6R_2)+S^{-(2)}AB(3n-8R_1+5R_2)$$

$$= S^{-4}S^{-(6)}A^3B^3[S^4\{S^3-3(n+4)\,S^2+(3n^2+42n+40)\,S-n(n^3+30n+176)\}$$

$$-kS^2\{3(n^2+2n-2)\,S^2-(3n^3+51n^2+72n-80)\,S+5n(n^2+6n-40)\}$$

$$-2k^2\{(n^3+9n^2+14n-12)\,S^2-3(n-2)\,(n+5)\,S-2n(n^2-4)\}$$

$$+32R_1\{6S^4-kS^2(S-1)\,(S+10)+k^2(S-1)^2\,(S+4)\}$$

$$-R_1\{2S^4(9S-128)-kS^2(3S^3-43S^2-168S-120)-2k^2(11S^3+23S^2+10S-4)\}$$

$$-R_2\,S\{120S^3-30kS^2(S+3)+k^2(S+1)\,(S^2+15S-4)\}].$$

But $$\overline{\chi^4} = S^2\left(1-\frac{2S\bar{x}}{AB}-\frac{S^2\overline{x^2}}{A^2B^2}\right),\quad \overline{\chi^6} = S^3\left(1-\frac{3S\bar{x}}{AB}+\frac{3S^2\overline{x^2}}{A^2B^2}-\frac{S^3\overline{x^3}}{A^3B^3}\right)$$

Hence $\overline{\chi^2} = S(S-1)^{-1}\,(n-1),$

$$\overline{\chi^4} = S(S-1)^{-(3)}[(n^2-1)\,S^2+6(2n-1)\,S-6R_1S^2-k\{-(n^2+2n-2)\,S$$

$$+n(n-2)+R_1S(S+1)\}],$$

$$\overline{\chi^6} = S(S-1)^{-(5)}[S^2\{(n-1)\,(n+1)\,(n+3)\,S^2+2(30n^2+69n-43)\,S+120(3n-1)\}$$

$$-kS\{(3n^3+21n^2+24n-26)\,S^2-(5n^3-57n^2-266n+120)\,S-60n(n-2)\}$$

$$+2k^2\{(n^3+9n^2+14n-12)\,S^2-3n(n-2)\,(n+5)\,S+2n(n^2-4)\}$$

$$-3nR_1\{6S^4-kS^2(S-1)\,(S+10)+k^2(S-1)^2\,(S+4)\}$$

$$-R_1\{2S^3(47S+180)-kS^2(19S^2+201S+180)+2k^2(11S^3+23S^2+10S-8)\}$$

$$+R_2\,S\{120S^3-30kS^2(S+3)+k^2(S+1)\,(S^2+15S-4)\}]. \tag{1}$$

These equations are rarely the most convenient. It is usually better to subtract the classical values, and write

$$\begin{aligned}
\overline{\chi^2} &= n-1+(S-1)^{-1}(n-1),\\
\overline{\chi^4} &= n^2-1+(S-1)^{-(3)}[(k-6)\,S\{R_1S(S+1)-(n^2+2n-2)\,S+n(n-2)\}\\
&\quad +6R_1S^2-(5n^2+12n-11)\,S+6(n^2-1)],\\
\overline{\chi^6} &= (n-1)(n+1)(n+3)+(S-1)^{-(5)}[S^3\{(5-k)(3n^3+21n^2+24n-26)\,S\\
&\quad +(3n-1)(S+120)\}-(n-1)(n+1)(n+3)(85S^3-225S^2+274S-120)\\
&\quad +kS^2\{(5n^3-57n^2-266n+120)\,S-60n(n-2)\}\\
&\quad +2k^2S\{(n^3+9n^2+14n-12)\,S^2-3n(n-2)(n+5)\,S+2n(n^2-4)\}\\
&\quad -3nR_1S\{6S^4-kS^2(S-1)(S+10)+k^2(S-1)^2(S+4)\}\\
&\quad -R_1S\{2S^3(47S+180)-kS^2(19S^2+210S+180)+2k^2(11S^3+23S^2+10S-8)\}\\
&\quad +R_2S^2\{120S^3-30kS^2(S+3)+k^2(S+1)(S^2+15S-4)\}]. \qquad (2)
\end{aligned}$$

To calculate the moments about the mean, we write

$$\overline{\chi^2}=n-1+\alpha,\quad \overline{\chi^4}=n^2-1+\beta,\quad \overline{\chi^6}=(n-1)(n+1)(n+3)+\gamma,$$

whence $\kappa_2=\mu_2=2(n-1)-2(n-1)\,\alpha+\beta-\alpha^2$,

$$\kappa_3=\mu_3=8(n-1)+3(n-1)(n-3)\,\alpha-3(n-1)(\beta-2\alpha^2)+\gamma-3\alpha\beta+2\alpha^3. \qquad (3)$$

The full algebraical expressions for these moments are rather cumbrous. However, by expanding equations (2) in descending powers of S, we have

$$\begin{aligned}
\kappa_1 &= \mu_1' = n-1+(n-1)\,S^{-1}+\dots,\\
\kappa_2 &= \mu_2 = 2(n-1)+(k-6)\,R_1-[k(n^2+2n-2)-2(2n^2+8n-7)]\,S^{-1}+\dots,\\
\kappa_3 &= \mu_3 = 8(n-1)+2(11k-56)\,R_1+(k^2-30k+120)\,R_2\\
&\quad -2(3n-2)[(3n+8)\,k-4(3n+11)]\,S^{-1}+\dots. \qquad (4)
\end{aligned}$$

In these equations the comparatively small terms involving R_1 and R_2 are omitted in the coefficients of S. The easiest forms for computation are

$$\begin{aligned}
\kappa_2 &= 2(n-1)+(k-6)\,R_1-[(k-4)(n^2+2n-2)-2(4n+1)]\,S^{-1},\\
\kappa_3 &= 8(n-1)+2[11(k-5)-1]\,R_1+[(k-15)^2-105]\,R_2\\
&\quad -2(3n-2)[(k-4)(3n+8)-12]\,S^{-1}.
\end{aligned}$$

These equations are sufficiently accurate for most practical applications.

The terms in equations (4) not involving negative powers of S are identical with the moments of χ^2 for $n-1$ degrees of freedom, used as a test of goodness of fit, derivable from Haldane's (1937) equations (3), if the expectations of a_r and b_r are ps_r and qs_r, and $pq=k^{-1}$. The effect of using χ^2 as a test of homogeneity rather than goodness of fit is to reduce the variance and the positive skewness, the reduction being considerable if k is large. There can also be little doubt, by analogy with the simpler case, that the fourth cumulant of χ^2 as a test of homogeneity is

$$\begin{aligned}
\kappa_4 &= 48(n-1)+96(4k-19)\,R_1+16(7k^2-125k+420)\,R_2\\
&\quad +(k^3-126k^2+1680k-5040)\,R_3+O(S^{-1}).
\end{aligned}$$

It will be seen that when $k-4$ is small, that is to say, the two types are almost equally frequent, both variance and skewness are reduced if samples are small. But when one class is fairly rare, so that k is large, they may be considerably increased.

In any particular case, one of three things may happen. It may be clear from equations (4) that the value of χ^2 is not significantly above that expected on a basis of homogeneity. This will certainly be so if $\chi^2-(n-1)$ is less than $\kappa_2^{\frac{1}{2}}$, and in this case there is no need to calculate κ_3. It may be clear that the corrections to the moments are unimportant, in which case the ordinary tables may be used. This is especially frequent for large values of n, where the distribution of χ^2 tends to normality. Or finally it is necessary to take κ_3 into consideration. If so we calculate κ_2 and κ_3 from equations (2), and make the transformation (Haldane, 1938)

$$\xi=\left[\left(\frac{\chi^2}{\kappa_1}\right)^h+\frac{h(1-h)\kappa_1}{2\kappa_1^2}-1\right]\frac{\kappa_1}{h\kappa_2^{\frac{1}{2}}}, \tag{5}$$

where $h=1-\dfrac{\kappa_1\kappa_3}{3\kappa_2^2}$. Then ξ is almost normally distributed with mean zero and variance unity. The probability that χ^2 should exceed its observed value is the probability that ξ should do so.

MOMENTS OF χ^2 WHEN ALL SAMPLES ARE EQUAL

If every s_r is equal, then $SR_1=n^2$, $S^2R_2=n^3$. So equations (1) become

$$\begin{aligned}
\overline{\chi^2}&=(n-1)\,S(S-1)^{-1},\\
\overline{\chi^4}&=(n-1)\,S(S-1)^{-(3)}[(n-1)\,S^2-6(n-1)\,S-2k(S-n)],\\
\overline{\chi^6}&=(n-1)\,(S-1)^{-(5)}[S^3\{(n+1)\,(n+3)\,S^2-2(9n^2+26n-43)\,S+120(n-1)^2\}\\
&\quad-2kS^2(S-n)\{(n+13)\,S-60(n-1)\}-4k^2(S-n)\{(n-6)\,S^2+9nS-4n\}],\\
\mu_2&=2(n-1)\,S(S-n)\,(S-1)^{-(3)}\left(\frac{S^2}{S-1}-k\right),\\
\mu_3&=2(n-1)\,(S-n)\,(S-1)^{-(5)}[4S^4(S-1)^{-2}\{S^2-(4n-5)\,S-2n\}\\
&\quad-kS^3(S-1)^{-1}\{(n-8)\,S-17n+10\}-4k^2\{(n-6)\,S^2+9nS-4n\}].
\end{aligned} \tag{6}$$

These equations, and those of the following paper, are a useful check on the accuracy of equations (1). They show that when the numbers in the samples are equal, or nearly so, the rarity of one class will always reduce μ_2, and will reduce μ_3 unless n is less than 8. However, μ_3 is always positive.

A NUMERICAL APPLICATION

Spurway (1945) investigated the frequency of crossing-over between a number of sex-linked genes in *Drosophila subobscura*, and used χ^2 as a test of the homogeneity of different families. In most cases no expectation was less than 5, and the classical method was used. However, for 9 pairs of genes the expectations were often very small, usually because crossing-over was infrequent. Her results are tabulated in Table 1. In the first column C denotes that the genes were in coupling, R that they were in repulsion, n is the number of families, S the total flies in them, A the number of flies showing crossing-over. Table 2 shows the actual data for the genes *bg* and *ct*, tabulated in the last row.

κ_2 was calculated from formulae (2) and (3), so the values are exact. It would have been quite sufficient to use formula (4). For example, in the case of *wi* and *y* this gives 234·9 instead of 235·8, and in that of *bg ct*, 17·570 in place of 17·507. The values of κ_2 are considerably greater than the classical value $2(n-1)$ in some cases, slightly less in others. κ_3 is calculated from formula (4) throughout. It may also be increased up to about 12 times, or slightly decreased. The accurate formula is not required, since, when n is large, a comparatively large change in κ_3 does not greatly affect P.

The values of P are calculated from equation (5) except when P is nearly 0·5, when the distribution is taken as normal. It will be seen that none of the values of χ^2 are significantly larger than their expectation, nor is their total. If no correction had been made, but the usual distribution of χ^2 had been employed, the values of P for *wi y*, *cp sn R*, and the total, would have been 0·035, 0·06 and 0·04 respectively. Thus the data would have been judged significantly heterogeneous.

Table 1. *Data on crossing-over in* Drosophila subobscura. *Explanation in text*

Genes	n	χ^2	S	A	R_1	R_2	κ_1	κ_2	κ_3	P
wi y C	63	83·245	1342	18	4·7266	0·56584	62·047	235·838	5915	0·10
cv l C	25	24·563	776	35	1·4129	0·12783	24·031	56·034	454·7	0·47
cv l R	66	64·662	1656	88	4·2961	0·49543	65·040	149·245	1106	0·51
l sn C	39	47·259	1167	386	2·2898	0·28976	38·033	72·162	266·1	0·14
l sn R	22	20·410	694	283	1·2712	0·12315	21·030	39·764	143·3	0·54
cp sn C	40	39·959	2359	113	1·1336	0·06469	39·017	83·467	499·8	0·46
cp sn R	103	121·970	5617	300	2·9840	0·17773	102·018	214·916	1220	0·09
sn bg C	10	6·523	273	45	0·8752	0·16451	9·033	17·993	83·43	0·25
bg ct C	10	9·231	273	54	0·8752	0·16451	9·033	17·507	74·85	0·48
	378	417·82					369·28	886·87	9763	0·07

Table 2. *Numbers of male flies, s_r, and cross-overs, a_r, between the loci of bg and ct, in ten families.* $\Sigma s_r^{-1} = 0{\cdot}8752$, $\Sigma s_r^{-2} = 0{\cdot}1645$. *Smallest expectation of a_r is* 0·593

s_r	23	68	28	12	28	7	7	38	3	59	273
a_r	2	18	3	2	7	0	1	10	1	10	54

Summary

χ^2 can be used as a test of homogeneity, even when expectations are less than unity, by the use of the formulae here given. In many cases the approximate formulae (4) are quite sufficient.

I have to thank Dr Spurway for the use of her numerical data, and for help with the computations.

References

Fisher, R. A. (1941). *Statistical Methods for Research Workers.* Edinburgh.

Haldane, J. B. S. (1937). The exact value of the moments of the distribution of χ^2, used as a test of goodness of fit when expectations are small. *Biometrika*, **29**, 389–91.

Haldane, J. B. S. (1938). The approximate normalization of a class of frequency distributions. *Biometrika*, **29**, 392–404.

Haldane, J. B. S. (1940). The mean and variance of χ^2, when used as a test of homogeneity, when expectations are small. *Biometrika*, **31**, 346–55.

Spurway, Helen (1945). The genetics and cytology of *Drosophila subobscura*. 1. Element A. Sex-linked mutants and their standard order. *J. Genet.* **46**, 268–86.

THE TREATMENT OF TIES IN RANKING PROBLEMS

By M. G. KENDALL

1. When a number of objects are presented for ranking by an observer there sometimes arise cases in which he is unable to express a preference in regard to certain of them and 'ranks them equal' or regards them as 'tying'. The effect may arise either because the objects really are indistinguishable, so far as the quality under consideration is concerned, or because the observer is unable to discern such differences as exist. Ties of this character are by no means uncommon—and indeed may be more the rule than the exception in some classes of work—and it is desirable to have a systematic method of dealing with them. In this paper I consider the effect of ties on coefficients of rank correlation, the estimation of rankings and the measurement of concordance in judges.

Rank correlations

2. The method of allocating ranking numbers to tied individuals in general use is to average the ranks which they cover. For instance, if the observer ties the third and fourth members each is allotted the number $3\frac{1}{2}$, and if he ties the second to the seventh inclusive, each is allotted the number $\frac{1}{6}(2+3+4+5+6+7) = 4\frac{1}{2}$. This is known as the mid-rank method and is the only one I shall consider. In fact I have seen only two other courses mentioned:

(a) 'Student' (1921) refers to a suggestion by Karl Pearson, as an alternative to mid-ranks, that the ties should all be ranked as if they were the highest member of the tie.

This is subject to the obvious disadvantages that it gives different results if one ranks from the other end of the scale and that it destroys the useful property that the mean rank of the whole series shall be $\frac{1}{2}(n+1)$. So far as I know it has never been used in problems involving the calculation of ranking coefficients.

(b) According to Woodbury (1940), DuBois (1939), in a paper which I have been unable to consult, has suggested allotting the ties an equal rank but proposes to determine it so that the sum of squares of the ranks shall be that of an untied ranking, namely, of the first n integers, $\frac{1}{6}n(n+1)(2n+1)$. This is rather troublesome, and it is not clear to me what advantages it possesses over the mid-rank method.

3. The effect of ties on the calculation of Spearman's ρ was considered by 'Student' (1921). ρ may be regarded as the product-moment correlation between two variates given by the two rankings. Since the variance of a ranking of n is given by $\frac{1}{12}(n^2-1)$, ρ is given by

$$\rho = \frac{12}{n^3-n}\sum_{i=1}^{n}\{X_i-\tfrac{1}{2}(n+1)\}\{Y_i-\tfrac{1}{2}(n+1)\}, \tag{1}$$

where X_i and Y_i represent the two rankings. This is easily reduced to the simpler and more familiar form

$$\rho = 1-\frac{6\Sigma(d^2)}{n^3-n}, \tag{2}$$

where $$d_i = X_i - Y_i. \tag{3}$$

Pursuing this analogy with the product-moment correlation 'Student' shows that, on the mid-rank method, the effect of a tie of t consecutive members is to lower the variance of the ranking by $\frac{1}{12n}(t^3-t)$. This is additive for any number of sets of ties in either the X- or the Y-variate, and if we write

$$T_X = \tfrac{1}{12}\Sigma(t^3-t), \tag{4}$$

the summation being over the ties of the X-variate, and T_Y for the similar sum for Y, we find for the product-moment correlation, say ρ_S,

$$\rho_S = \frac{1}{2}\frac{\operatorname{var} X + \operatorname{var} Y - \operatorname{var}(X-Y)}{\sqrt{(\operatorname{var} X \operatorname{var} Y)}}$$

$$= \frac{\frac{1}{6}(n^3-n)-(T_X+T_Y)-\Sigma(d^2)}{\sqrt{\{\frac{1}{6}(n^3-n)-2T_X\}\{\frac{1}{6}(n^3-n)-2T_Y\}}} \tag{5}$$

$$= \frac{\frac{1}{6}(n^3-n)-(T_X+T_Y)-\Sigma(d^2)}{\{\frac{1}{6}(n^3-n)-(T_X+T_Y)\}\sqrt{\left[1-\frac{(T_X-T_Y)^2}{\{\frac{1}{6}(n^3-n)-(T_X+T_Y)\}^2}\right]}}. \tag{6}$$

'Student' notes that if $T_X - T_Y$ is small, we have approximately

$$\rho_S = 1 - \frac{\Sigma(d^2)}{\frac{1}{6}(n^3-n)-(T_X+T_Y)}. \tag{7}$$

It is also useful to note that if T_X and T_Y are small compared with $\frac{1}{6}(n^3-n)$, we have

$$\rho_S = 1 - \frac{6\Sigma(d^2)}{n^3-n},$$

so that the correction to be applied to the ordinary formula is negligible for many practical purposes.

Example 1. Consider, for instance, the two rankings of 10:

X:	1	$2\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{1}{2}$	$6\frac{1}{2}$	8	$9\frac{1}{2}$	$9\frac{1}{2}$
Y:	1	2	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	8	8	8	10

In the first ranking there are four tied pairs and hence

$$T_X = \tfrac{4}{12}(2^3-2) = 2.$$

In the second there is one set of four ties and one of three, and hence

$$T_Y = \tfrac{1}{12}(60+24) = 7.$$

We also have $\Sigma(d^2) = 13.$

Hence, in accordance with (5), $\rho_S = \frac{165-22}{\sqrt{(161.151)}} = \frac{143}{155\cdot 92}$

$= 0\cdot 9171.$

Calculation on the basis of (2) gives $\rho = 1 - \frac{78}{990} = 0\cdot 9212.$

The value given by (7) is $\rho_S = 1 - \frac{13}{165-9}$

$= 0\cdot 9167.$

4. There is another way of looking at this problem which 'Student' did not mention. Suppose we regard any set t of tied ranks as due to inability on the part of the observer to distinguish real differences; i.e. we assume that there does exist a set of integral ranks although we are ignorant of it on present evidence. Then we may ask, what is the *average* value of ρ over all the $t!$ possible ways of assigning integral ranks to the tied members?

5. If the t corresponding members in the Y-ranking are held fixed, then the average covariance for all $t!$ arrangements of the X-members is the covariance of the fixed Y-members and the average of the t X-members. But this latter gives the mean ranks of the tied members, and consequently the mean covariance of the two rankings is

$$\frac{1}{n}\{\tfrac{1}{12}(n^3-n)-\tfrac{1}{2}\Sigma(d^2)-\tfrac{1}{2}(T_X+T_Y)\}, \tag{8}$$

the effect of various sets of ties being additive. If now we divide by the *actual* variances of X and Y we arrive at equation (5). Thus 'Student's' formula may be regarded as giving a mean value of the coefficient which would be obtained if the ties were replaced in all possible ways by the integral ranks which they cover; always bearing in mind that we have not averaged the variances.

6. A similar point of view has been adopted by Woodbury (1940) who does not seem to have been aware of 'Student's' results; but Woodbury takes as his variance the quantity $\frac{1}{12}(n^2-1)$, that is to say, he determines the average ρ which would be obtained if the ties were replaced by appropriate integral rankings in all possible ways, the variance in each case being that of the first n integers. This results in ρ_W, say, where

$$\rho_W = 1-\frac{6\{\Sigma(d^2)+T_X+T_Y\}}{n^3-n}, \tag{9}$$

the difference from ρ_S of equation (5) lying, of course, in the denominators in the second term on the right.

Example 2. For example, in the illustration considered above Woodbury's value would be

$$\rho_W = 1-\frac{6(13+2+7)}{990}$$
$$= 0{\cdot}8667,$$

against $\rho_S = 0{\cdot}9171$.

7. The question then arises, which is the better measure of rank correlation, ρ_S or ρ_W? It is useful in the first instance to consider some special cases.

(*a*) Suppose that the two rankings are both completely tied, i.e. that each rank is $\frac{1}{2}(n+1)$. We clearly have

$$\rho = 1,$$

indicating complete correlation. For the 'Student' form we have

$$\rho_S = \frac{0}{\sqrt{(0\times 0)}},$$

an indeterminate form which, however, may be regarded as unity as a limiting case in virtue of the next subsection. For Woodbury's form we find

$$\rho_W = \frac{0}{\frac{1}{6}(n^3-n)}$$
$$= 0,$$

indicating zero correlation. In short, Woodbury's 'correction' has reduced ρ from 1 to 0.

(*b*) Suppose that both rankings are the same, that the last member in each is ranked n and that the others are all tied and hence have rank $\frac{1}{2}n$. Then it will be found that

$$\rho = 1,$$
$$\rho_S = 1,$$
$$\rho_W = \frac{3}{n+1} \quad (n \geqslant 2).$$

The crude form of ρ and 'Student's' corrected form are in agreement that the correlation is unity. Woodbury's form differs entirely and gives a correlation which is small for large n.

(*c*) Generally, if the two rankings are identical and there are ties giving a T-number of T in each

$$\rho = \rho_S = 1,$$
$$\rho_W = 1 - \frac{12T}{n^3 - n}.$$

(*d*) Suppose that one ranking is in the natural order $1, \ldots, n$ and has no ties, so that $T_X = 0$. If the other ranking has the last member ranked n and the others completely tied we find

$$\Sigma(d^2) = \tfrac{1}{12}n(n-1)(n-2),$$
$$\rho = \frac{n+4}{2(n+1)} \sim \frac{1}{2},$$
$$\rho_S = \sqrt{\frac{3}{n+1}},$$
$$\rho_W = \frac{3}{n+1}.$$

For large n, ρ tends to 0·5, whereas ρ_S and ρ_W tend to zero, the latter faster than the former.

8. It appears to me that the decision as to which of the coefficients ρ_S or ρ_W is preferable can only rest on the use to which they are to be put.

(*a*) Let us suppose in the first instance that the objects have a definite ranking $1, \ldots, n$ determined in some objective way. The purpose of correlating the order assigned by an observer is then to determine the observer's accuracy, not the real ranking. 'Student's' form of the coefficient would measure the product-moment correlation of ranks, giving weight to the fact that if the observer produces ties the variance of his estimates is reduced. Woodbury's form would measure the average correlation of all the results obtained if the observer allotted to the tied groups integral ranks determined at random. For instance in a ranking of 8

1	2	3	4	5	6	7	8
4	4	4	4	4	4	4	8

$\rho_S = \sqrt{\frac{1}{3}} = 0{\cdot}577$, $\rho_W = \frac{1}{3} = 0{\cdot}333$. There does not seem to me to be much to choose between the two, but on the whole Woodbury's form gives figures nearer to what one would expect. We may suppose the observer to be in a genuine state of indecision when considering the tied members, and the average of all the values given by guessing integral ranks at random seems a fair measure of his ability. 'Student's' form gives higher values because he divides the product-moment by the actual standard deviations, and hence gives the observer credit, so to speak, for clustering his values in spite of the fact that he ought not to do so *because there really is an objective order*. In a case of this kind I should therefore favour Woodbury's form.

(b) The situation is quite different if no objective order is given and we are measuring the concordance between two judges. In this case Woodbury's form seems to me to give the wrong answer. In the case where two rankings are identical, for instance, one is entitled to expect that a measure of correlation should be unity—agreement could not be better. Both judges may be wrong, but that is not the point. We are measuring their agreement, not their accuracy. It has been shown above that if all members are tied Woodbury's form would give a zero correlation between the two rankings, which on the face of it seems ridiculous. We are no longer entitled to assume an objective order, or, even if there are real differences in the objects, to suppose that they fall above the threshold of the discriminatory power of the judges. 'Student's' form appears to be far better.

9. It is, of course, undeniable that 'Student's' form is more troublesome to calculate. This is unimportant if only two or three rankings are to be compared, but might be more important if there were large numbers of rankings. In such a case, however, it is more usual to work out a single measure of joint correlation rather than many pairs. The problem of m rankings for tied variates is dealt with below.

10. I proceed to consider the appropriate method of dealing with ties in calculating the alternative coefficient of correlation known as τ (Kendall, 1938). In an elegant synthesis of the rank correlation problem Daniels (1944) has recently pointed out that τ may be defined as

$$\tau = \frac{\Sigma(a_{ij} b_{ij})}{\sqrt{\{\Sigma a_{ij}^2 \Sigma b_{ij}^2\}}}, \tag{10}$$

where $$a_{ij} = -a_{ji}, \quad b_{ij} = -b_{ji},$$

a_{ij} is a score allotted to each pair of ranks X_i, X_j as $+1$ if $j > i$ and -1 if $i < j$, b_{ij} similarly relating to the Y-ranking. In the ordinary ranking case, of course,

$$\Sigma a_{ij}^2 = \Sigma b_{ij}^2 = \tfrac{1}{2}n(n-1). \tag{11}$$

Daniels' form has the advantage of revealing τ as analogous to an ordinary product-moment coefficient.

11. To extend this definition to the case of tied ranks we have only to define the score a_{ij} for equal ranks, and this is easily done by defining it as zero, midway between the scores of $+1$ and -1 taken when the ranks are unequal. This, it will be noticed, affects the denominator in the definition of equation (10) as well as the numerator.

Example 3. Let us consider the rankings of Example 1, namely,

X:	1	$2\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{1}{2}$	$6\frac{1}{2}$	8	$9\frac{1}{2}$	$9\frac{1}{2}$
Y:	1	2	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	8	8	8	10

Considering the first member in association with the other 9, we see that the score in both rankings in each case is $+1$, so that the total score is 9; the second and third members in the X-ranking are tied and this pair therefore scores 0, whatever the Y-position. The score from pairs associated with the second member will be found to be 7; in the Y-ranking members 3–6 are tied and therefore the only non-vanishing scores arise from association of the third member with the seventh, eighth, ninth and tenth members, score 4. The total score will be found to be

$$9+7+4+4+4+3+1+1+0 = 33.$$

The sum Σa_{ij}^2 is found to be 41 and Σb_{ij}^2 is 36 and hence, writing τ_s for the corresponding quantity to ρ_s, we have

$$\tau_s = \frac{33}{\sqrt{(41.36)}} = 0{\cdot}8589.$$

The value of ρ_S is 0·9171 but the difference need cause no concern as the coefficients do give rather different results, having different scales.

The general rule for the formation of Σa_{ij}^2 will be clear. If there is a tie of extent t we calculate $\frac{1}{2}t(t-1)$ and sum for all ties. If this sum is U then

$$\Sigma a_{ij}^2 = \tfrac{1}{2}n(n-1) - U. \tag{12}$$

12. If we replace any tied set by the corresponding integral ranks in any order and average for all the $t!$ possible orders we get the same result as by replacing a_{ij} for the tied members by zero; for in the $t!$ arrangement each pair will occur an equal number of times in the order XY and the order YX, so that the allocation of $+1$ in the first case and -1 in the second is equivalent to the allocation of zero on the average. Thus we may regard our score for tied ranks as the mean of the values obtained by allotting integral ranks in all possible ways. On the analogy of Woodbury's treatment of ρ we could then define

$$\tau_W = \frac{\text{score}}{\frac{1}{2}n(n-1)}. \tag{13}$$

The choice between τ_S and τ_W is precisely the same as in the case of Spearman's ρ; that is to say we might use the latter where an objective order is known but the former where it is a question of measuring the concordance between judges.

13. For the purposes of comparison with the special cases considered in § 7 it may be worth while giving the corresponding values of T:

(*a*) Both rankings completely tied:

$$\tau_S \text{ indeterminate}, \quad \tau_W = 0.$$

(*b*) Rankings identical and all tied except last member:

$$\tau_S = 1, \quad \tau_W = \frac{2}{n}.$$

(*c*) Rankings identical, ties giving U-member U in each:

$$\tau_S = 1, \quad \tau_W = 1 - \frac{2U}{n(n-1)}.$$

(*d*) One ranking equal to the natural order $1, \ldots, n$, the other all tied except the last member:

$$\tau_S = \sqrt{\frac{2}{n}}, \quad \tau_W = \frac{2}{n}.$$

The estimation of a ranking

14. In a previous note (1942) I considered the problem of estimating the true ranking (or the ranking on which there was the greatest measure of agreement) for m-rankings of n individual exemplified by

Object	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
	4	2	1	7	6	3	5	8
	7	2	1	6	4	5	3	8
	7	4	2	6	5	3	1	8
Sum of ranks	18	8	4	19	15	11	9	24

It was shown that a reasonable estimate was to be obtained by ranking according to the sums of ranks, beginning with the lowest, e.g. in this example the ranking would be

$$A_3 \quad A_2 \quad A_7 \quad A_6 \quad A_5 \quad A_1 \quad A_4 \quad A_8.$$

This is the best in that it minimizes the sum of squares of deviations from what they would be if the m-rankings were identical; and it also maximizes the average ρ between the observed and the estimated rankings.

15. The above method may also be regarded in this way: if an object is ranked r, it is preferred to $n-r$ members but $r-1$ members are preferred to it. Allotting as usual $+1$ for the first type and -1 to the second we see that the individual scores $n+1-2r$ in its own ranking. Summing over the m-rankings we see that an individual ranked X_r, Y_r, Z_r, etc. has altogether a score of

$$m(n+1)-2\Sigma(X). \tag{14}$$

If then we rank the individuals according to their total scores, beginning with the highest, we arrive at exactly the same result as by ranking according to $\Sigma(X)$ beginning with the lowest. Thus our method arranges the objects in the order of numbers of preferences; a further argument in its favour. It is also easy to see that the method minimizes the sums of squares of deviations of preferences from what they would be if there were complete agreement. In fact, denote the estimated ranking by $X_1, \ldots, X_n$ and let the corresponding sums of preferences be $\xi_1, \xi_2, \ldots, \xi_n$, this being a permutation of $m(n+1)-mX_j$, $j = 1, \ldots, n$. If the actual preferences are given in sum by $S_1, \ldots, S_n$ we have to minimize

$$\sum_{j=1}^{n} (S_j-\xi_j)^2 = \Sigma S^2+\Sigma\xi^2-2\Sigma(S\xi). \tag{15}$$

The first two terms on the right are constants and we have therefore to maximize $\Sigma(S\xi)$. This is clearly done by multiplying the largest S by the largest ξ, that is to say the largest S by the smallest X, and so on. In other words, we allot X_1 to the largest S and so on in order.

16. To complete the story one would like to be able to prove that the method maximized the average τ between observed and estimated rankings.

Unfortunately the proposition fails, as is shown by the following example:

A_1	A_2	A_3	A_4
2	3	1	4
1	2	4	3
1	2	4	3
4	7	9	10

The estimated order here would be that running from left to right across the page as written, and the total score between that order and the three observed orders will be found to be $2+4+4 = 10$. But if we interchange the last two columns it becomes $0+6+6 = 12$. Such a situation, however, is of rare occurrence and can only occur when there is substantial disagreement between judges on the two objects interchanged, in which case no ranking is very reliable. I do not think, therefore, that the failure of the result in extreme cases is important.

17. Suppose now that some of the ranks are tied. Does the method of summing ranks apply unchanged to give a good estimate?

(*a*) In the first place, the method continues to give an answer which appears reasonable on the face of it. Moreover it may be regarded as an average result for all the ways of permuting the tied ranks when replaced by the appropriate integral ranks.

(*b*) If the question is regarded as one of ranking according to preferences, the replacement of a pair of integral ranks by a tie does not affect the preferences with other members and

merely cancels a preference between the tied pair; and so for any set of ties. In consequence the method preserves the property of ranking according to the number of preferences.

(c) If we measure the average ρ with the estimated ranking in Woodbury's form of corrected ρ the method provides a maximum average ρ unless the estimated ranking itself contains ties, in which case the result might conceivably fail, though it is unlikely to do so.

In fact, let the estimated ranking be $X_1, \ldots, X_n$ and the rank of the jth object in the kth ranking be Y_{jk}. We shall maximize the average ρ by maximizing

$$V = \sum_{k=1}^{m} \sum_{j=1}^{n} \{X_j - \tfrac{1}{2}(n+1)\}\{Y_{jk} - \tfrac{1}{2}(n+1)\}$$

$$= \sum_{j=1}^{n} \{X_i - \tfrac{1}{2}(n+1)\}\{S_j - \tfrac{1}{2}m(n+1)\}, \tag{16}$$

which reduces to maximizing $\Sigma(XS)$. If, however, there are ties in the estimated ranking our problem is to minimize something of the form

$$\frac{V}{\left\{\tfrac{1}{2}(n^2-1) - \frac{1}{n}\Sigma(T_X)\right\}^{\frac{1}{2}}}, \tag{17}$$

and variations in T_X may upset the result. This, however, is not likely to be serious unless there are many ties in the estimated ranking, in which case estimation of any kind is unreliable.

(*d*) If we measure the average ρ with the estimated ranking in 'Student's' form the result again may fail for (16) then becomes of the form

$$\Sigma\Sigma\{X_j - \tfrac{1}{2}(n+1)\} A_{jk}\{Y_{jk} - \tfrac{1}{2}(n+1)\}, \tag{18}$$

where the coefficients A_{jk} differ from our ranking to another because they depend on differing variances.

(*e*) Similar considerations apply to the proposition that the method minimizes the sums of squares of deviations from what the sums of ranks or preferences would be if all rankings were alike. Apart from complications due to ties in the estimated ranking, the minimal properties continue to obtain.

18. To sum up, therefore, the method of estimating the ranking according to sums of ranks appears to give satisfactory results when ties are involved.

Example 4. Consider the three rankings

	1	2	3	$4\frac{1}{2}$	$4\frac{1}{2}$	6	$7\frac{1}{2}$	$7\frac{1}{2}$	9	10
	1	$2\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{1}{2}$	$6\frac{1}{2}$	8	$9\frac{1}{2}$	$9\frac{1}{2}$
	1	2	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	8	8	8	10
Sums of ranks	3	$6\frac{1}{2}$	10	$13\frac{1}{2}$	$13\frac{1}{2}$	17	22	$23\frac{1}{2}$	$26\frac{1}{2}$	$29\frac{1}{2}$
Estimated ranking	1	2	3	$4\frac{1}{2}$	$4\frac{1}{2}$	6	7	8	9	10

The sums of ranks give an estimated ranking as shown. There is one case here where the sums of ranks are equal and the individual ranks yielding those sums are also equal. There seems no better course than to tie them. Had the sums for these two been

$4\frac{1}{2}$	$3\frac{1}{2}$
$4\frac{1}{2}$	$4\frac{1}{2}$
$4\frac{1}{2}$	$5\frac{1}{2}$
$13\frac{1}{2}$	$13\frac{1}{2}$

we might perhaps have ranked the former as 4 and the latter as 5, on the ground that the variance of the former is less and the group therefore 'cluster' better than the other. This is a new principle deriving no support from the various minimal principles already introduced. It is the usual practice, I think, to regard an estimate as better when it is based on more closely grouped observations; but here the resulting estimate of the mean ranking is $4\frac{1}{2}$ so it can also be argued that the tie should remain. On the whole this seems to me the better course.*

TESTS OF SIGNIFICANCE FOR m RANKINGS

19. I proceed to consider what modifications, if any, are required in significance tests when ties can appear in the rankings. Babington Smith & I (1939) have discussed the case when the ranks are integral. The algebra required for the more extended discussion has been given by Pitman (1938) in considering a similar problem in the analysis of variance.

Consider an array of m rows

$$\begin{array}{cccc} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ \dots & \dots & \dots & \dots \\ k_1 & k_2 & \dots & k_n \end{array} \tag{19}$$

If S is the sum of square of column totals about their mean and S' the sum of squares of all values about their mean, we define

$$W = \frac{S}{mS'} \tag{20}$$

as the coefficient of concordance. In the case when the a's, b's, etc. are permutations of the natural numbers 1 to n we have

$$W = \frac{S}{\frac{1}{12}m^2(n^3-n)}. \tag{21}$$

W can vary from 0 to 1, attaining the latter value only if all rankings are identical.

20. Let us in the first place consider what happens to formula (21) when some of the integral ranks are replaced by ties. If the T-numbers for the various rows are T_a, T_b, etc., the formula becomes

$$W = \frac{S}{\frac{1}{12}m^2(n^3-n) - m\Sigma(T)}. \tag{22}$$

This is as simple a form as we require.

In the data of Example 4, for instance, we find

$$S = 682{\cdot}9,$$

$$W = \frac{682{\cdot}9}{742{\cdot}5 - 30}$$

$$= 0{\cdot}958.$$

* In my note (1942), which dealt only with integral ranks, I suggested that, where the sums of ranks are equal, precedence should be given to the one with the smaller variance; but I was there considering only an estimated ranking which itself was integral. When ties are permitted I should, as stated above, use them where sums of ranks are equal.

21. Denoting by α_r the rth moment of the a-row in (19), and similarly for β_r, γ_r, etc., and by α'_r the rth k-statistic, we have for the moments of W, from Pitman's results,

$$\overline{W} = E(W) = \frac{1}{m}, \tag{23}$$

$$E(W-\overline{W})^2 = \frac{4}{m^2(n-1)}\frac{\Sigma\alpha_2\beta^2}{\Sigma^2\alpha_2}, \tag{24}$$

$$E(W-\overline{W})^3 = \frac{48}{m^3(n-1)}\frac{\Sigma\alpha_2\beta_2\gamma_2}{\Sigma^3\alpha_2}+\frac{8(n-1)(n-2)}{m^3n^4}\frac{\Sigma\alpha'_3\beta'_3}{\Sigma^3\alpha_2}, \tag{25}$$

$$E(W-\overline{W})^4 = \frac{48}{m^4(n-1)^2}\frac{\Sigma^2\alpha_2\beta_2}{\Sigma^4\alpha_2}-\frac{96}{m^4(n-1)^2(n+1)}\frac{\Sigma\alpha_2^2\beta_2^2}{\Sigma^4\alpha_2}+\frac{1152}{m^4(n-1)^3}\frac{\Sigma\alpha_2\beta_2\gamma_2\delta_2}{\Sigma^4\alpha_2}$$
$$+\frac{16(n-1)(n-2)(n-3)}{m^4n^5(n+1)}\frac{\Sigma\alpha'_4\beta'_4}{\Sigma^4\alpha_2}+\frac{252(n-2)}{m^4n^4}\frac{\Sigma\alpha'_3\beta'_3\gamma_2}{\Sigma^4\alpha_2}. \tag{26}$$

In the case when the numbers are permutations of the first n integers these expressions reduce to

$$\overline{W} = \frac{1}{m}, \tag{27}$$

$$E(W-\overline{W})^2 = \frac{2(m-1)}{m^3(n-1)}, \tag{28}$$

$$E(W-\overline{W})^3 = \frac{8(m-1)(m-2)}{m^5(n-1)^2}, \tag{29}$$

$$E(W-\overline{W})^4 = \frac{12(m-1)^2}{m^6(n-1)^2}+\frac{48(m-1)(m-2)(m-3)}{m^7(n-1)^3}-\frac{48(m-1)}{m^7(n-1)^2(n+1)}. \tag{30}$$

If m and n are moderately large, these expressions are approximately the same (exactly so for the first two) as the moments of

$$dF = \frac{1}{B(p,q)}W^{p-1}(1-W)^{q-1}dW, \tag{31}$$

where

$$\left.\begin{aligned} p &= \tfrac{1}{2}(n-1)-\frac{1}{m},\\ q &= (m-1)p. \end{aligned}\right\} \tag{32}$$

It follows that W can be tested in Fisher's z-distribution by writing

$$z = \tfrac{1}{2}\log\frac{(m-1)W}{1-W}, \tag{33}$$

$$\left.\begin{aligned} \nu_1 &= (n-1)-\frac{2}{m},\\ \nu_2 &= (m-1)\nu_1. \end{aligned}\right\} \tag{34}$$

How far does this require modification for tied ranks?

22. For the purposes of an accurate test we can, of course, work out the first four moments of W from (23) to (26) in individual cases and fit an *ad hoc* curve; but this is a tedious process and some approximation is necessary.

The first two moments of (31) are

$$\mu_1 \text{ (about zero)} = \frac{p}{p+q},$$

$$\mu_2 = \frac{pq}{(p+q)^2(p+q+1)},$$

and if we identify them with the first two moments of W, $\frac{1}{m}$ and $\mu_2(W)$, say, we find

$$\left.\begin{aligned} p &= -\frac{1}{m}+\frac{m-1}{m^3\mu_2(W)}, \\ q &= (m-1)p, \end{aligned}\right\} \tag{35}$$

so that approximately W can be tested in the z-distribution with

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+\frac{2(m-1)}{m^3\mu_2(W)}, \\ \nu_2 &= (m-1)\,\nu_1. \end{aligned}\right\} \tag{36}$$

23. We have, as in (24),

$$\mu_2(W) = \frac{4}{m^2(n-1)}\frac{\Sigma\alpha_2\beta_2}{\Sigma^2\alpha_2}.$$

Writing A for $\Sigma\alpha_2$ and B for $\Sigma\alpha_2^2$ we have

$$\mu_2(W) = \frac{2}{m^2(n-1)}\left\{1-\frac{B}{A^2}\right\},$$

so that the appropriate degrees of freedom are

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+\frac{(n-1)(m-1)}{m(1-B/A^2)}, \\ \nu_2 &= (m-1)\,\nu_1. \end{aligned}\right\} \tag{37}$$

If the T-numbers appropriate to the various rankings are small compared with $\frac{1}{12}(n^3-n)$ we can approximate further. In fact write N for $\frac{1}{12}(n^3-n)$. Then

$$A = mN-\Sigma(T),$$

$$B = mN^2-2N\Sigma(T)+\Sigma(T^2),$$

and to the first order in $\Sigma(T)$

$$\begin{aligned} \frac{B}{A^2} &= \frac{mN^2-2N\Sigma(T)}{m^2N^2-2mN\Sigma(T)} \\ &= \frac{1}{m}\left\{1-\frac{2\Sigma(T)}{mN}\right\}\left\{1-\frac{2\Sigma(T)}{mN}\right\}^{-1} \\ &= \frac{1}{m}. \end{aligned}$$

On substitution in (27)

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+(n-1), \\ \nu_2 &= (m-1)\,\nu_1. \end{aligned}\right\} \tag{38}$$

This, of course, is the same as (34) so that, if the number or extent of the ties is small, the test for untied ranks requires no modification (other than that necessary in the calculation of W itself).

24. It thus appears that we can apply the usual test unchanged unless the ties are numerous enough to render $\Sigma(T)$ not small compared with mN. If the ties are numerous we can work with the modified degrees of freedom given by (35), but in such a case it would probably be as well to verify by direct calculation that the third and fourth moments of W were in reasonable agreement with those given by the β-approximation implicit in the use of the z-test. If it happens that one or two rankings contribute the major part of $\Sigma(T)$ we may perhaps reject them on the grounds that the judges are very bad, but the rejection of observations has to be done with some care and only after we are satisfied that they really are exceptional and not merely outlying members of a continuous range.

Addition of extra members to a ranking

25. There is one class of case in which I have found the coefficient τ to have definite advantages over ρ. An example will illustrate the point best. Suppose I send out an inquiry to a number of firms asking for some information which they may or may not wish to disclose; and suppose that the information is of a type for which one would expect that the non-participants might differ from the participants. By a certain date a number of replies have been received and it is then necessary to close the inquiry and to summarize the results. How far can I assume that the replies to hand are representative of the population to which the inquiry was addressed? Is there any evidence to suggest that those who reply earlier to the inquiry differ from those who reply later?

26. To simplify the illustration suppose that I receive 15 replies in the form of a percentage figure which occur in the following order:

Order of receipt:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Percentage:	15	13	12	16	25	8	9	14	17	11	18	20	10	21	19
Rank of percentage:	8	6	5	9	15	1	2	7	10	4	11	13	3	14	12

I have chosen percentages which are all different so as not to complicate the example, but if ties are present there is no essential modification.

Now if there is some relation between the order of reply and the magnitude of the percentage, it ought to be shown up by the rank correlation between the order of reply and the order of magnitude of the percentage. The latter is shown in the last row of the example above and we find

$$\Sigma(d^2) = 392,$$

$$\rho = 0{\cdot}300.$$

This in fact, is barely significant, but I am not for the moment concerned with significance. Suppose that after we have completed this calculation two more replies arrive with percentages 7 and 23. We now have to calculate a revised value of ρ by re-numbering nearly all the replies and working *ab initio*. In practice the continual arrival of stragglers is quite common and to work out a new value of ρ each time is a great arithmetical nuisance. The point I wish to make is that τ is not subject to this disability, extra values being capable of addition as required.

In the above example for 15 members the value of $\Sigma(a_{ij}b_{ij})$ is easily seen to be

$$0+3+4+1-10+9+8+3+2+3+2+1+2-1 = 27,$$

so that

$$\tau = \tfrac{27}{105} = 0{\cdot}257.$$

If now we add a 16th member with the value 7, the contribution to $\Sigma(ab)$ is obtained by considering this new member in conjunction with the other fifteen, and is seen to be -15. Similarly, a further member valued 23 adds 13. The new value of $\Sigma(ab)$ is thus 25 and the new τ is given by

$$\tau = \tfrac{25}{136} = 0{\cdot}184.$$

In this way a kind of running value of τ can be ascertained without re-ranking at each stage as is necessary with ρ. Thus τ has a decided advantage in this class of case, namely the calculation of ranking coefficients for time series which may be extended in length.

REFERENCES

DANIELS, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika*, **33**, 129.

DUBOIS, P. (1939). Formulas and tables for rank correlation. *Psychol. Rec.* **3**, 45.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81.

KENDALL, M. G. (1942). Note on the estimation of a ranking. *J.R. Statist. Soc.* **105**, 119.

KENDALL, M. G. & BABINGTON SMITH, B. (1939). The problem of *m* rankings. *Ann. Math. Statist.* **10**, 275.

PITMAN, E. J. G. (1938). Significance tests which may be applied to samples from any population. III. The analysis of variance test. *Biometrika*, **29**, 322.

'STUDENT' (1921). An experimental determination of the probable error of Dr Spearman's correlation coefficient. *Biometrika*, **13**, 263.

WOODBURY, M. A. (1940). Rank correlation when there are equal variates. *Ann. Math. Statist.* **11**, 358.

THE PROBABILITY INTEGRAL OF THE MEAN DEVIATION

EDITORIAL NOTE

1. About 3 years ago a need arose to obtain the probability levels of the mean deviation in random samples from a normal population. The requirement was in a field of production quality control where it was customary, as well as convenient, to use the mean deviation as a measure of dispersion in a sample, rather than the standard deviation or the range between extreme individuals. The need was pressing, and it appeared that the quickest answer for practical purposes would be obtained by using the known expressions for the mean and variance of the M.D. and getting a measure of the departure of the distributions from normality by a sampling experiment with random numbers. An investigation on these lines was undertaken by Dr E. H. Sealy and Mr C. D. Bates of the Advisory Service on Quality Control, Ministry of Supply. Their results, in the form of a table of factors for control limits for sample sizes varying from $n = 5$ to 20, were issued in 1943.

2. These limits, though adequate for the immediate requirement, were of course not exact and are now superseded in the range $n = 2$ to 10 by the tables printed below in Mr Godwin's paper.

3. At the time when the planning of the earlier investigation was discussed with Dr Sealy, I had overlooked the fact that R. C. Geary's paper of 1936 on the distribution of the ratio of the M.D. to the S.D. in samples from a normal population contained formulae from which the higher moments of the M.D. could be derived. When this oversight was realized and before Mr Godwin's work was undertaken, I had asked Dr Geary to develop from his earlier work, expansions for the 3rd and 4th order semi-invariants of the M.D. in terms of inverse powers of $\nu = n-1$ (where n is the sample size). It seems of interest to take this opportunity of putting these results on record.

4. *Dr Geary's expansions* (population standard deviation as unit).
Mean deviation in a sample of n observations:

$$m = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|. \tag{1}$$

Expectation of m:

$$\bar{m} = \sqrt{\left\{\frac{2(n-1)}{\pi n}\right\}}. \tag{2}$$

Semi-invariants of the sampling distribution of m when the population is normal:

$$\lambda_3 = \left(\frac{\nu}{\nu+1}\right)^{\frac{3}{2}} \left\{\frac{0\cdot218\,014}{\nu^2} - \frac{0\cdot074\,170}{\nu^3} + \frac{0\cdot057\,313}{\nu^4} - \frac{0\cdot040\,457}{\nu^5} + \frac{0\cdot023\,601}{\nu^6} - \dots\right\}, \tag{3}$$

$$\lambda_4 = \left(\frac{\nu}{\nu+1}\right)^{2} \left\{\frac{0\cdot114\,771}{\nu^3} - \frac{0\cdot068\,509}{\nu^4} + \frac{0\cdot033\,371}{\nu^5} - \frac{0\cdot120\,003}{\nu^6} + \dots\right\}. \tag{4}$$

The frequency constants β_1 and β_2 can be obtained from

$$\beta_1 = \lambda_3^2/\lambda_2^3, \quad \beta_2 = 3 + \lambda_4/\lambda_2^2, \tag{5}$$

where

$$\lambda_2 = \sigma_m^2 = \frac{2(n-1)}{n^2\pi}\left\{\tfrac{1}{2}\pi + \sqrt{\{n(n-2)\}} - n + \sin^{-1}\frac{1}{n-1}\right\}. \tag{6}$$

It should be noted that Geary (1936) has taken n' for the sample size and written $n = n' - 1$, but to be consistent with the paper which follows, I have written n for his n' and ν for his $n = n' - 1$.*

5. The accompanying table shows the 3rd and 4th semi-invariants of m computed from the expansions given above and also the moment ratios β_1 and β_2; for $n = 4$ the figures for β_1 and β_2 may be compared with R. A. Fisher's (1920) values of $\beta_1 = 0{\cdot}297$, $\beta_2 = 3{\cdot}280$. Differences between the expansion and true values will become rapidly less as n increases.

The sampling moments of the mean deviation

Sample size n	λ_3	λ_4	β_1	β_2
4	0·014 32	0·001 900	0·299	3·244
5	0·009 057	0·000 9785	0·230	3·194
6	0·006 244	0·000 5636	0·187	3·160
8	0·003 483	0·000 2351	0·136	3·118
10	0·002 218	0·000 1193	0·106	3·093
12	0·001 536	0·000 06864	0·088	3·076
15	0·000 9800	0·000 03492	0·069	3·060
20	0·000 5495	0·000 01461	0·051	3·045

REFERENCES

Fisher, R. A. (1920). *Mon. Not. R. Astr. Soc.* **80**, 8.

Geary, R. C. (1936). *Biometrika*, **28**, 295.

Sealy, E. H. & Bates, C. D. (1943). Ministry of Supply, Advisory Service on Quality Control. *Technical Report*, QC/R/1.

* In equations (8) and (22) of Geary's paper (pp. 296 and 300), for d read $d' = \sum_{i=1}^{n'} |x_i - \bar{x}|/\sqrt{\{n'(n'-1)\}}$.

E.S.P

ON THE DISTRIBUTION OF THE ESTIMATE OF MEAN DEVIATION OBTAINED FROM SAMPLES FROM A NORMAL POPULATION

BY H. J. GODWIN, *Advisory Service on Statistical Method, Ministry of Supply*

The relative merits and demerits of the mean deviation and standard deviation as measures of the dispersion of a population have been discussed by Fisher (1920): though the balance is rather in favour of the latter, the mean deviation is widely used, especially in experimental work where many small samples are taken and where saving in computation is a consideration. The distribution of the estimate given by random samples from a Normal population has not previously been obtained, save for the special cases of sample sizes four (by Fisher, 1920) and five (by Jones): Helmert (1876), and later Fisher, found the second moment of the distribution, and Geary (1936) found the third and fourth moments and showed that β_1 and β_2 were $O(n^{-1})$ and $O(n^{-1})$ respectively for large sample size n. Thus the distribution may be approximated to by a Normal distribution, and this approximation improves as n increases. The distribution was estimated empirically, for small sample sizes, by Sealy & Bates (1943). In the present paper an expression for the distribution for general n is found, suitable for calculation by quadratures and Table 1 gives the resulting probability integral to 5 decimal places for $n = 2$ to 10. From this table certain percentage points have been calculated and are given in Table 2. The Normal approximations to the percentage points for sample size ten are given for comparison with the true values: Sealy's values were much closer, being least good for the extreme percentage points of the smallest sample sizes.

Let the sample values be $x_1, x_2, \ldots, x_n$, where the x's are distributed according to the frequency function $\frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2} dx$ (there is no loss of generality in taking the population mean and standard deviation to be zero and unity respectively). The numbers $1, 2, \ldots, n$ can be assigned to the members of the sample in $n!$ ways: we suppose that $x_1 \leqslant x_2 \leqslant x_3 \ldots \leqslant x_n$. We consider separately the cases when the mean falls between x_1 and x_2, x_2 and x_3, ..., x_{n-1} and x_n. Suppose $x_k \leqslant \frac{\Sigma x_i}{n} \leqslant x_{k+1}$.

Then the mean deviation

$$m = \frac{(x_{k+1} + \ldots + x_n) - (x_1 + \ldots + x_k) - \frac{n-2k}{n} \Sigma x_i}{n},$$

i.e.
$$\frac{n^2 m}{2} = k(x_{k+1} + \ldots + x_n) - (n-k)(x_1 + \ldots + x_k).$$

The frequency function of m is found by evaluating

$$(2\pi)^{-\frac{1}{2}n} \int \ldots \int e^{-\frac{1}{2}(\Sigma x_i^2)} dx_1 \ldots dx_n \tag{1}$$

over the region defined by
$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n, \tag{2}$$

$$x_k \leqslant \frac{\Sigma x_i}{n} \leqslant x_{k+1}, \tag{3}$$

and
$$\frac{n^2 m}{2} \leqslant k(x_{k+1} + \ldots + x_n) - (n-k)(x_1 + \ldots + x_k) \leqslant \frac{n^2(m+dm)}{2}. \tag{4}$$

The various functions so obtained are summed over k from 1 to $n-1$, and the whole multiplied by $n!$. The integral is evaluated by a transformation of the quadratic form Σx_i^2: this is most easily done in two stages.

First put $$x_{i+1}-x_i = y_i.$$

Then (2) becomes $$y_i \geqslant 0, \tag{5}$$

(3) becomes and

$$\left.\begin{aligned} &y_1+2y_2+\ldots+(k-1)\,y_{k-1} \leqslant (n-k)\,y_k+\ldots+y_{n-1}\\ &y_1+2y_2+\ldots+ky_k \leqslant (n-k-1)\,y_{k+1}+\ldots+y_{n-1}\end{aligned}\right\} \tag{6}$$

and (4) becomes

$$\frac{n^2m}{2} \leqslant (n-k)\,y_1+2(n-k)\,y_2+\ldots+k(n-k)\,y_k+k(n-k-1)\,y_{k+1}+\ldots+ky_{n-1} \leqslant \frac{n^2(m+dm)}{2}. \tag{7}$$

Now put $$u_j = y_1+2y_2+\ldots+jy_j \quad (j \leqslant k-1),$$
$$u_j = (n-j-1)\,y_{j+1}+\ldots+y_{n-1} \quad (j \geqslant k).$$

Then (5), (6) and (7) become

$$\left.\begin{aligned} &O \leqslant u_1 \leqslant u_2 \leqslant \ldots \leqslant u_{k-1} \leqslant \frac{n(m+dm)}{2},\\ \text{and}\quad &O \leqslant u_{n-2} \leqslant u_{n-3} \leqslant \ldots \leqslant u_k \leqslant \frac{n(m+dm)}{2},\end{aligned}\right\} \tag{8}$$

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(x_1, m, n, \ldots, u_{n-2})} = \frac{n^2}{2k!(n-k)!},$$

and $$\sum x_i^2 = nx_1^2+\sum_{j=1}^{n-1} 2x_1(n-j)\,y_j+\sum_{j=1}^{n-1}(n-j)\,y_j^2+2\sum_{j=1}^{n-2}\sum_{l=j+1}^{n-1}(n-l)\,y_jy_l$$
$$= n\left(x_1+\sum_{j=1}^{k-1}\frac{u_j}{j(j+1)}+\frac{nm}{2k}\right)^2+\sum_1^{k-1}\frac{u_j^2}{j(j+1)}+\sum_k^{n-2}\frac{u_j^2}{(n-j)(n-j-1)}+\frac{m^2n^3}{4k(n-k)}.$$

We now define a series of functions $G_r(x)$, such that

$$G_0(x) = 1, \quad G_r(x) = \int_0^x \exp\left[-\frac{t^2}{2r(r+1)}\right] G_{r-1}(t)\,dt. \tag{9}$$

The integral (1) now appears as the product of n simple integrals and, after integration subject to the restrictions (8), gives

$$\frac{n^2}{2k!(n-k)!}(2\pi)^{-\frac{1}{2}n}\frac{\sqrt{(2\pi)}}{\sqrt{n}}\exp\left[-\frac{m^2n^3}{8k(n-k)}\right] G_{k-1}\left(\frac{nm}{2}\right) G_{n-k-1}\left(\frac{nm}{2}\right) dm.$$

The frequency function of m is accordingly

$$f_n(m)\,dm = n!\sum_{k=1}^{n-1}\frac{n^2}{2k!(n-k)!}(2\pi)^{-\frac{1}{2}n}\frac{\sqrt{(2\pi)}}{\sqrt{n}}\exp\left[-\frac{m^2n^3}{8k(n-k)}\right] G_{k-1}\left(\frac{nm}{2}\right) G_{n-k-1}\left(\frac{nm}{2}\right) dm$$
$$= \frac{n^{\frac{3}{2}}}{2^{\frac{1}{2}(n+1)}\pi^{\frac{1}{2}(n-1)}}\left\{\sum_{k=1}^{n-1} {}^nC_k \exp\left[-\frac{m^2n^3}{8k(n-k)}\right] G_{k-1}\left(\frac{nm}{2}\right) G_{n-k-1}\left(\frac{nm}{2}\right)\right\} dm. \tag{10}$$

The calculation of the G-functions, the distribution function of m, and the percentage points was done under the direction of Dr H. O. Hartley, whose care and assistance I gratefully acknowledge: a note by him on the method of computation appears below as an Appendix.

Although the theoretical part of the work was done privately, the computations were carried out, and the work prepared for publication, under the auspices of the Ministry of Supply (S.R. 17) from whom permission to publish has been obtained.

REFERENCES

FISHER, R. A. (1920). *Mon. Not. R. Astr. Soc.* 80, 8, June 1920.

GEARY, R. C. (1936). *Biometrika*, 28, 295–305.

HELMERT, R. F. (1876). *Astr. Nachr.* 88, no. 2096.

JONES, A. E. Unpublished paper.

SEALY, E. H. & BATES, C. D. (1943). Ministry of Supply, Advisory Service on Quality Control. *Technical Report*, QC/R/1.

APPENDIX

NOTE ON THE CALCULATION OF THE DISTRIBUTION OF THE ESTIMATE OF MEAN DEVIATION IN NORMAL SAMPLES

BY H. O. HARTLEY, *Scientific Computing Service, Ltd.*

The formula used for the computation of this distribution function is the finite series (10) of the above paper. This expression involves the functions $G_r(x)$ which are defined by the recurrence formula (9). The numerical work therefore consists of:

(*a*) The calculation of the $G_r(x)$ by a recurrence of numerical quadratures.

(*b*) The calculation of the distribution functions $f_n(m)$ from formula (10).

(*c*) The numerical quadratures $\int_0^m f_n(m)\,dm$ yielding, for each n, the probability integral for the mean deviation m.

(*a*) The essential feature of the numerical quadratures is a new method on the National Accounting Machine. With this method, of which it is hoped to publish details in due course, it is possible to produce a table of the integral $\int_a^x f(x)\,dx$ from the 4th differences of the integrand $f(x)$ in a single operation.

The starting point of the recurrence was the table of

$$\frac{1}{\sqrt{\pi}}G_1(x) = \frac{1}{\sqrt{\pi}}\int_0^x e^{-\frac{1}{4}t^2}\,dt = \frac{2}{\sqrt{\pi}}\int_0^{\frac{1}{2}x} e^{-\alpha^2}\,d\alpha,$$

given in *Tables of the Probability Integral*, vol. **1**, W.P.A., New York.

This integral was multiplied by the ordinate

$$\frac{2}{\sqrt{\pi}}e^{-\frac{1}{4}x^2} = \frac{2}{\sqrt{\pi}}e^{-(\frac{1}{2}x)^2},$$

which is tabulated next to it in the W.P.A. Tables. Products were formed at interval 0·05 in x and these were differenced. The function*

$$(480)\,\pi^{-1}\,2^{-\frac{1}{2}}\,G_2(x)$$

was then obtained by the mechanical method of numerical quadrature in accordance with formula (9).

This process was then repeated for $r = 3, \ldots, 8$; producing the functions $G_r(x)$ with increasing constant factors and for ranges as shown below:

r	Factors of $G_r(x)$	Range
2	$480\ \pi^{-1}\,2^{-\frac{1}{2}}$	$x = 0\ (0{\cdot}05)\ 13{\cdot}5$
3	$480^2\,\pi^{-\frac{3}{2}}\,2^{-1}$	$x = 0\ (0{\cdot}05)\ 19{\cdot}0$
4	$480^3\,\pi^{-2}\,2^{-\frac{3}{2}}$	$x = 0\ (0{\cdot}05)\ 25{\cdot}0$
5	$480^4\,\pi^{-\frac{5}{2}}\,2^{-3}$	$x = 0\ (0{\cdot}1)\ 31{\cdot}0$
6	$480^5\,\pi^{-3}\,2^{-\frac{9}{2}}$	$x = 0\ (0{\cdot}1)\ 36{\cdot}0$
7	$480^6\,\pi^{-\frac{7}{2}}\,2^{-6}$	$x = 0\ (0{\cdot}1)\ 39{\cdot}0$
8	$480^7\,\pi^{-4}\,2^{-\frac{15}{2}}$	$x = 0\ (0{\cdot}1)\ 46{\cdot}0$

* The factor (480) is necessitated by the method of quadrature. The same applies to those shown in the table below.

Seven significant figures were accurate in the maximum value of G_2, but this accuracy gradually decreased to 5 significant figures for the maximum value of G_8. The ordinate functions

$$\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{t^2}{2r(r+1)}\right]$$

which occur as multipliers in formula (9) were obtained by interpolation in the Tables of z (Table II, *Tables for Statisticians and Biometricians*, vol. **1**).

(*b*) For convenience of computation formula (10) was rewritten as follows:

$$f_n(m) = n^{\frac{1}{2}}\, 2^{-\frac{1}{2}(n+1)}\, \pi^{-\frac{1}{2}(n-1)} \sum_{k=1}^{n} {}^nC_k\, g_{k-1}(x)\, g_{n-k-1}(x), \tag{11}$$

where $x = \frac{1}{2}nm$ and $g_r(x) = G_r(x)\exp\left[-\dfrac{x^2}{2(r+1)}\right]$.

Using the symmetry in k the number of terms may be halved.

The G-functions were first converted to g-functions through multiplication by the ordinates

$$\frac{c_r}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2(r+1)}\right] \quad (c_r = \text{suitably chosen constant}).$$

These ordinates were obtained by interpolation in the z-tables of Table II of *Tables for Statisticians and Biometricians*, vol. **1**. The x-interval was 0·05 for $g_0, \ldots, g_4$ and 0·2 for $g_5, \ldots, g_8$. Formula (11) was then applied to obtain $f_n(m)$ at the following x- and m-intervals:

r	3	4	5	6	7	8	9	10
x-interval	0·075	0·10	0·125	0·15	0·175	0·2	0·225	0·25
m-interval	0·05	0·05	0·05	0·05	0·05	0·05	0·05	0·05

In certain cases it will be seen that the x-interval is not a tabular interval for the g-functions. In such cases Lagrangian Interpolation had to be applied. Finally the $f_n(m)$ functions were subtabulated to the final m-interval of 0·01.

(*c*) The published tables of the probability integrals $\int_0^m f_n(m)\,dm$ were then obtained by the process of mechanical quadrature on the National Machine. The tables of percentage points were obtained by inverse interpolation.

Checks were as follows. Apart from the usual checks by differencing, the first and second moments μ_1' and μ_2' were calculated from the final tables of $f_n(m)$ as well as $\int_0^m f_n(m)\,dm$. These were compared with check values calculated from the theoretical formulae

$$\mu_1' = \left(\frac{2(n-1)}{n\pi}\right)^{\frac{1}{2}},$$

$$\mu_2' = \frac{2(n-1)}{n^2\pi}\left\{2\tan^{-1}\left(\frac{n}{n-2}\right)^{\frac{1}{2}} + n^{\frac{1}{2}}(n-2)^{\frac{1}{2}}\right\}.$$

Five-decimal agreement was obtained throughout.

Grateful acknowledgement is made to Mr M. Sumner for the expert help rendered in the calculation of these tables.

TABLES OF THE

PROBABILITY INTEGRAL OF THE

MEAN DEVIATION IN NORMAL SAMPLES

Table 1. *The probability integral of the mean deviation* (m) *in normal samples of n observations.* (*Population standard deviation as unit*)

m \ n	2	3	4	5	6	7	8	9	10
0·00	0·00000	0·00000							
·01	·01128	·00019	0·00000						
·02	·02256	·00074	·00003	0·00000					
·03	·03384	·00167	·00009	·00001					
·04	·04511	·00297	·00022	·00002					
0·05	0·05637	0·00464	0·00042	0·00004	0·00000				
·06	·06762	·00668	·00073	·00009	·00001				
·07	·07886	·00908	·00115	·00016	·00002	0·00000			
·08	·09008	·01184	·00172	·00027	·00004	·00001			
·09	·10128	·01496	·00244	·00042	·00007	·00001			
0·10	0·11246	0·01843	0·00333	0·00064	0·00012	0·00002	0·00000		
·11	·12362	·02226	·00442	·00093	·00019	·00004	·00001		
·12	·13476	·02644	·00571	·00130	·00030	·00007	·00002	0·00000	
·13	·14587	·03095	·00723	·00178	·00044	·00011	·00003	·00001	
·14	·15695	·03581	·00899	·00237	·00064	·00017	·00005	·00001	0·00000
0·15	0·16800	0·04100	0·01101	0·00310	0·00089	0·00026	0·00008	0·00002	0·00001
·16	·17901	·04651	·01329	·00398	·00122	·00037	·00012	·00004	·00001
·17	·18999	·05234	·01585	·00503	·00163	·00053	·00018	·00006	·00002
·18	·20094	·05849	·01871	·00626	·00214	·00074	·00026	·00009	·00003
·19	·21184	·06495	·02187	·00770	·00277	·00101	·00038	·00014	·00005
0·20	0·22270	0·07171	0·02534	0·00936	0·00354	0·00135	0·00053	0·00021	0·00008
·21	·23352	·07876	·02914	·01126	·00445	·00179	·00073	·00030	·00012
·22	·24430	·08610	·03327	·01342	·00554	·00232	·00099	·00042	·00018
·23	·25502	·09371	·03773	·01585	·00682	·00297	·00132	·00059	·00026
·24	·26570	·10160	·04254	·01858	·00830	·00377	·00173	·00080	·00037
0·25	0·27633	0·10974	0·04769	0·02161	0·01002	0·00472	0·00225	0·00108	0·00052
·26	·28690	·11814	·05320	·02497	·01199	·00585	·00289	·00143	·00072
·27	·29742	·12679	·05907	·02867	·01423	·00717	·00366	·00188	·00097
·28	·30788	·13567	·06528	·03271	·01677	·00873	·00459	·00244	·00130
·29	·31828	·14478	·07186	·03713	·01962	·01052	·00571	·00312	·00172
0·30	0·32863	0·15410	0·07879	0·04192	0·02280	0·01259	0·00703	0·00395	0·00224
·31	·33891	·16364	·08608	·04709	·02634	·01495	·00858	·00496	·00289
·32	·34913	·17337	·09372	·05266	·03025	·01763	·01039	·00616	·00368
·33	·35928	·18330	·10171	·05864	·03455	·02065	·01248	·00759	·00465
·34	·36936	·19340	·11005	·06503	·03926	·02404	·01488	·00928	·00582
0·35	0·37938	0·20367	0·11872	0·07183	0·04439	0·02783	0·01762	0·01124	0·00722
·36	·38933	·21410	·12773	·07905	·04996	·03202	·02073	·01352	·00887
·37	·39921	·22469	·13706	·08670	·05598	·03666	·02424	·01615	·01082
·38	·40901	·23541	·14671	·09476	·06247	·04175	·02817	·01915	·01309
·39	·41874	·24626	·15667	·10325	·06942	·04731	·03256	·02257	·01573
0·40	0·42839	0·25724	0·16693	0·11215	0·07686	0·05338	0·03742	0·02643	0·01877
·41	·43797	·26832	·17748	·12147	·08478	·05995	·04279	·03076	·02224
·42	·44747	·27951	·18831	·13120	·09319	·06705	·04869	·03560	·02618
·43	·45689	·29079	·19941	·14133	·10209	·07468	·05513	·04099	·03063
·44	·46623	·30215	·21075	·15186	·11148	·08286	·06215	·04693	·03563
0·45	0·47548	0·31358	0·22234	0·16277	0·12136	0·09160	0·06975	0·05347	0·04121
·46	·48466	·32507	·23416	·17405	·13172	·10089	·07796	·06063	·04741
·47	·49375	·33661	·24620	·18569	·14255	·11074	·08677	·06843	·05425
·48	·50275	·34820	·25843	·19768	·15386	·12115	·09621	·07689	·06177
·49	·51167	·35982	·27086	·21001	·16562	·13212	·10627	·08602	·06998
0·50	0·52050	0·37146	0·28345	0·22265	0·17783	0·14364	0·11697	0·09584	0·07892

Table 1 (cont.). *The probability integral of the mean deviation* (m) *in normal samples of* n *observations.* (*Population standard deviation as unit*)

m \ n	2	3	4	5	6	7	8	9	10
0·50	0·52050	0·37146	0·28345	0·22265	0·17783	0·14364	0·11697	0·09584	0·07892
·51	·52924	·38313	·29620	·23559	·19047	·15569	·12829	·10635	·08860
·52	·53790	·39479	·30909	·24881	·20352	·16828	·14023	·11756	·09903
·53	·54646	·40646	·32211	·26230	·21697	·18137	·15279	·12947	·11022
·54	·55494	·41811	·33525	·27604	·23079	·19497	·16595	·14207	·12218
0·55	0·56332	0·42975	0·34847	0·28999	0·24497	0·20904	0·17970	0·15536	0·13491
·56	·57162	·44135	·36178	·30416	·25948	·22357	·19402	·16931	·14840
·57	·57982	·45293	·37516	·31851	·27430	·23852	·20889	·18392	·16264
·58	·58792	·46446	·38858	·33302	·28941	·25389	·22427	·19916	·17761
·59	·59594	·47593	·40204	·34768	·30477	·26963	·24016	·21501	·19329
0·60	0·60386	0·48735	0·41552	0·36245	0·32037	0·28572	0·25650	0·23144	0·20965
·61	·61168	·49871	·42901	·37733	·33617	·30213	·27328	·24840	·22667
·62	·61941	·50999	·44249	·39228	·35215	·31882	·29046	·26588	·24431
·63	·62705	·52120	·45594	·40729	·36827	·33576	·30799	·28383	·26252
·64	·63459	·53232	·46936	·42233	·38452	·35293	·32585	·30220	·28127
0·65	0·64203	0·54335	0·48273	0·43739	0·40087	0·37027	0·34398	0·32095	0·30050
·66	·64938	·55428	·49603	·45244	·41727	·38777	·36235	·34004	·32016
·67	·65663	·56511	·50926	·46746	·43372	·40537	·38092	·35941	·34021
·68	·66378	·57583	·52240	·48244	·45017	·42306	·39965	·37902	·36058
·69	·67084	·58644	·53544	·49734	·46661	·44078	·41848	·39882	·38122
0·70	0·67780	0·59693	0·54836	0·51217	0·48300	0·45852	0·43739	0·41875	0·40206
·71	·68467	·60729	·56117	·52689	·49932	·47623	·45632	·43877	·42306
·72	·69143	·61753	·57384	·54148	·51555	·49388	·47523	·45882	·44414
·73	·69810	·62764	·58637	·55594	·53166	·51143	·49409	·47886	·46525
·74	·70468	·63762	·59874	·57025	·54762	·52886	·51284	·49883	·48634
0·75	0·71116	0·64745	0·61096	0·58439	0·56342	0·54614	0·53147	0·51868	0·50735
·76	·71754	·65714	·62300	·59834	·57903	·56324	·54992	·53838	·52821
·77	·72382	·66669	·63487	·61210	·59443	·58012	·56816	·55788	·54888
·78	·73001	·67609	·64655	·62564	·60961	·59677	·58616	·57713	·56930
·79	·73610	·68534	·65804	·63897	·62454	·61316	·60388	·59609	·58943
0·80	0·74210	0·69443	0·66934	0·65206	0·63922	0·62926	0·62130	0·61474	0·60922
·81	·74800	·70337	·68043	·66492	·65362	·64506	·63839	·63303	·62864
·82	·75381	·71215	·69131	·67752	·66773	·66054	·65512	·65092	·64763
·83	·75952	·72078	·70198	·68986	·68154	·67567	·67147	·66840	·66617
·84	·76514	·72924	·71243	·70193	·69503	·69045	·68742	·68544	·68422
0·85	0·77067	0·73754	0·72266	0·71373	0·70821	0·70486	0·70295	0·70201	0·70175
·86	·77610	·74568	·73266	·72525	·72105	·71888	·71805	·71810	·71875
·87	·78144	·75366	·74244	·73649	·73355	·73252	·73270	·73368	·73519
·88	·78669	·76147	·75199	·74744	·74571	·74575	·74689	·74874	·75105
·89	·79184	·76912	·76132	·75810	·75752	·75857	·76061	·76327	·76632
0·90	0·79691	0·77660	0·77040	0·76847	0·76898	0·77097	0·77386	0·77727	0·78099
·91	·80188	·78392	·77926	·77854	·78008	·78296	·78663	·79072	·79505
·92	·80677	·79107	·78789	·78832	·79082	·79453	·79891	·80363	·80851
·93	·81156	·79806	·79628	·79780	·80120	·80567	·81071	·81599	·82135
·94	·81627	·80489	·80444	·80698	·81122	·81640	·82202	·82780	·83359
0·95	0·82089	0·81155	0·81237	0·81587	0·82089	0·82671	0·83286	0·83907	0·84522
·96	·82542	·81805	·82007	·82447	·83021	·83660	·84321	·84981	·85626
·97	·82987	·82439	·82754	·83277	·83917	·84608	·85310	·86001	·86671
·98	·83423	·83057	·83478	·84079	·84778	·85515	·86252	·86970	·87659
·99	·83851	·83659	·84180	·84852	·85605	·86382	·87149	·87887	·88591
1·00	0·84270	0·84245	0·84860	0·85597	0·86398	0·87210	0·88001	0·88755	0·89468

Table 1 (cont.). *The probability integral of the mean deviation (m) in normal samples of n observations. (Population standard deviation as unit)*

m \ n	2	3	4	5	6	7	8	9	10
1·00	0·84270	0·84245	0·84860	0·85597	0·86398	0·87210	0·88001	0·88755	0·89468
·01	·84681	·84816	·85518	·86315	·87158	·87999	·88809	·89574	·90292
·02	·85084	·85371	·86154	·87005	·87885	·88751	·89575	·90347	·91065
·03	·85478	·85911	·86768	·87668	·88581	·89465	·90299	·91074	·91788
·04	·85865	·86436	·87362	·88305	·89245	·90144	·90984	·91756	·92464
1·05	0·86244	0·86946	0·87935	0·88916	0·89878	0·90788	0·91629	0·92397	0·93095
·06	·86614	·87442	·88487	·89502	·90482	·91398	·92237	·92997	·93681
·07	·86977	·87922	·89020	·90063	·91057	·91976	·92809	·93557	·94227
·08	·87333	·88389	·89533	·90600	·91604	·92521	·93346	·94081	·94733
·09	·87680	·88841	·90027	·91114	·92123	·93037	·93850	·94568	·95202
1·10	0·88020	0·89280	0·90502	0·91605	0·92616	0·93523	0·94322	0·95022	0·95635
·11	·88353	·89705	·90959	·92074	·93084	·93980	·94764	·95444	·96034
·12	·88679	·90117	·91398	·92521	·93527	·94411	·95176	·95835	·96403
·13	·88997	·90516	·91820	·92947	·93947	·94815	·95561	·96198	·96741
·14	·89308	·90901	·92224	·93353	·94344	·95195	·95920	·96533	·97052
1·15	0·89612	0·91275	0·92613	0·93740	0·94718	0·95551	0·96253	0·96843	0·97337
·16	·89910	·91635	·92985	·94108	·95072	·95885	·96564	·97128	·97598
·17	·90200	·91984	·93341	·94458	·95406	·96197	·96851	·97391	·97836
·18	·90484	·92321	·93682	·94790	·95720	·96488	·97118	·97633	·98054
·19	·90761	·92647	·94009	·95105	·96016	·96760	·97365	·97855	·98252
1·20	0·91031	0·92961	0·94321	0·95404	0·96294	0·97014	0·97594	0·98058	0·98432
·21	·91296	·93264	·94620	·95687	·96555	·97250	·97805	·98245	·98595
·22	·91553	·93556	·94905	·95955	·96800	·97470	·97999	·98415	·98743
·23	·91805	·93838	·95177	·96209	·97030	·97675	·98178	·98571	·98877
·24	·92051	·94109	·95437	·96449	·97246	·97865	·98343	·98713	·98998
1·25	0·92290	0·94371	0·95685	0·96676	0·97448	0·98041	0·98495	0·98842	0·99107
·26	·92524	·94623	·95921	·96890	·97637	·98204	·98634	·98959	·99206
·27	·92751	·94865	·96146	·97092	·97813	·98355	·98761	·99066	·99294
·28	·92973	·95098	·96360	·97283	·97978	·98495	·98878	·99162	·99373
·29	·93190	·95322	·96564	·97462	·98132	·98624	·98985	·99250	·99445
1·30	0·93401	0·95538	0·96758	0·97631	0·98275	0·98743	0·99082	0·99329	0·99508
·31	·93606	·95745	·96942	·97790	·98408	·98852	·99172	·99401	·99566
·32	·93807	·95944	·97117	·97940	·98533	·98953	·99253	·99465	·99616
·33	·94002	·96135	·97283	·98081	·98648	·99046	·99327	·99523	·99662
·34	·94191	·96318	·97441	·98213	·98755	·99132	·99394	·99576	·99702
1·35	0·94376	0·96494	0·97591	0·98337	0·98855	0·99210	0·99455	0·99623	0·99738
·36	·94556	·96662	·97733	·98453	·98947	·99282	·99510	·99665	·99770
·37	·94731	·96824	·97867	·98562	·99033	·99348	·99560	·99703	·99798
·38	·94902	·96979	·97995	·98664	·99112	·99409	·99606	·99736	·99823
·39	·95067	·97127	·98115	·98759	·99186	·99464	·99647	·99767	·99845
1·40	0·95229	0·97269	0·98229	0·98849	0·99254	0·99515	0·99684	0·99794	0·99865
·41	·95385	·97405	·98337	·98932	·99316	·99561	·99717	·99818	·99882
·42	·95538	·97534	·98439	·99010	·99374	·99603	·99748	·99839	·99897
·43	·95686	·97659	·98535	·99083	·99427	·99641	·99775	·99858	·99910
·44	·95830	·97777	·98626	·99151	·99477	·99676	·99799	·99875	·99922
1·45	0·95970	0·97891	0·98712	0·99215	0·99522	0·99708	0·99821	0·99890	0·99932
·46	·96105	·97999	·98793	·99274	·99564	·99737	·99841	·99904	·99941
·47	·96237	·98103	·98869	·99329	·99602	·99763	·99859	·99915	·99949
·48	·96365	·98201	·98941	·99380	·99637	·99786	·99874	·99926	·99956
·49	·96490	·98296	·99009	·99427	·99669	·99808	·99889	·99935	·99962
1·50	0·96611	0·98385	0·99073	0·99472	0·99699	0·99828	0·99901	0·99943	0·99967

Table 1 (cont.). *The probability integral of the mean deviation* (m) *in normal samples of n observations.* (*Population standard deviation as unit*)

m \ n	2	3	4	5	6	7	8	9	10
1·50	0·96611	0·98385	0·99073	0·99472	0·99699	0·99828	0·99901	0·99943	0·99967
·51	·96728	·98471	·99133	·99513	·99726	·99845	·99913	·99950	·99972
·52	·96841	·98552	·99190	·99551	·99751	·99861	·99923	·99957	·99976
·53	·96952	·98630	·99243	·99586	·99774	·99875	·99932	·99962	·99979
·54	·97059	·98704	·99293	·99619	·99795	·99888	·99940	·99967	·99982
1·55	0·97162	0·98774	0·99340	0·99650	0·99814	0·99900	0·99947	0·99971	0·99984
·56	·97263	·98841	·99384	·99678	·99831	·99911	·99953	·99975	·99987
·57	·97360	·98905	·99425	·99704	·99847	·99920	·99959	·99978	·99989
·58	·97455	·98966	·99464	·99728	·99861	·99929	·99964	·99981	·99990
·59	·97546	·99023	·99500	·99750	·99875	·99936	·99968	·99984	·99992
1·60	0·97635	0·99078	0·99534	0·99771	0·99887	0·99943	0·99972	0·99986	0·99993
·61	·97721	·99130	·99566	·99790	·99898	·99949	·99975	·99988	·99994
·62	·97804	·99179	·99596	·99808	·99908	·99955	·99978	·99990	·99995
·63	·97884	·99226	·99624	·99824	·99917	·99960	·99981	·99991	·99996
·64	·97962	·99270	·99650	·99839	·99925	·99964	·99983	·99992	·99996
1·65	0·98038	0·99312	0·99675	0·99852	0·99932	0·99968	0·99986	0·99993	0·99997
·66	·98110	·99352	·99698	·99865	·99939	·99972	·99987	·99994	·99997
·67	·98181	·99390	·99719	·99877	·99945	·99975	·99989	·99995	·99998
·68	·98249	·99425	·99739	·99887	·99951	·99978	·99990	·99996	·99998
·69	·98315	·99459	·99758	·99897	·99956	·99980	·99992	·99996	·99998
1·70	0·98379	0·99491	0·99775	0·99906	0·99960	0·99983	0·99993	0·99997	0·99999
·71	·98441	·99521	·99791	·99915	·99964	·99985	·99994	·99997	·99999
·72	·98500	·99550	·99806	·99922	·99968	·99986	·99994	·99998	·99999
·73	·98558	·99577	·99820	·99929	·99971	·99988	·99995	·99998	·99999
·74	·98614	·99603	·99834	·99936	·99974	·99989	·99996	·99998	·99999
1·75	0·98667	0·99627	0·99846	0·99941	0·99977	0·99991	0·99996	0·99999	0·99999
·76	·98719	·99650	·99857	·99947	·99979	·99992	·99997	·99999	·99999
·77	·98769	·99671	·99868	·99952	·99982	·99993	·99997	·99999	·99999
·78	·98817	·99691	·99878	·99956	·99983	·99993	·99998	·99999	1·00000
·79	·98864	·99710	·99887	·99960	·99985	·99994	·99998	·99999	
1·80	0·98909	0·99729	0·99895	0·99964	0·99987	0·99995	0·99998	0·99999	
·81	·98952	·99746	·99903	·99967	·99988	·99995	·99998	·99999	
·82	·98994	·99762	·99911	·99970	·99990	·99996	·99999	1·00000	
·83	·99035	·99777	·99917	·99973	·99991	·99996	·99999		
·84	·99074	·99791	·99924	·99976	·99992	·99997	·99999		
1·85	0·99111	0·99804	0·99930	0·99978	0·99993	0·99997	0·99999		
·86	·99147	·99817	·99935	·99980	·99993	·99997	·99999		
·87	·99182	·99829	·99940	·99982	·99994	·99998	·99999		
·88	·99216	·99840	·99945	·99984	·99995	·99998	·99999		
·89	·99248	·99850	·99949	·99986	·99995	·99998	1·00000		
1·90	0·99279	0·99860	0·99953	0·99987	0·99996	0·99998			
·91	·99309	·99869	·99957	·99988	·99996	·99999			
·92	·99338	·99878	·99960	·99989	·99997	·99999			
·93	·99366	·99886	·99963	·99990	·99997	·99999			
·94	·99392	·99894	·99966	·99991	·99998	·99999			
1·95	0·99418	0·99901	0·99969	0·99992	0·99998	0·99999			
·96	·99443	·99907	·99972	·99993	·99998	·99999			
·97	·99466	·99914	·99974	·99993	·99998	1·00000			
·98	·99489	·99920	·99976	·99994	·99998				
·99	·99511	·99925	·99978	·99994	·99999				
2·00	0·99532	0·99930	0·99980	0·99995	0·99999				

Table 1 (cont.). *The probability integral of the mean deviatian* (m) *in normal samples of n observations.* (*Population standard deviation as unit*)

m \ n	2	3	4	5	6
2·00	0·99532	0·99930	0·99980	0·99995	0·99999
·01	·99552	·99935	·99981	·99995	·99999
·02	·99572	·99940	·99983	·99996	·99999
·03	·99591	·99944	·99984	·99996	·99999
·04	·99609	·99948	·99986	·99997	·99999
2·05	0·99626	0·99951	0·99987	0·99997	0·99999
·06	·99642	·99955	·99988	·99998	·99999
·07	·99658	·99958	·99989	·99998	1·00000
·08	·99673	·99961	·99990	·99998	
·09	·99688	·99964	·99991	·99998	
2·10	0·99702	0·99966	0·99992	0·99998	
·11	·99716	·99969	·99992	·99999	
·12	·99728	·99971	·99993	·99999	
·13	·99741	·99973	·99994	·99999	
·14	·99753	·99975	·99994	·99999	
2·15	0·99764	0·99977	0·99995	0·99999	
·16	·99775	·99979	·99995	·99999	
·17	·99785	·99980	·99996	·99999	
·18	·99795	·99982	·99996	·99999	
·19	·99805	·99983	·99996	·99999	
2·20	0·99814	0·99984	0·99997	0·99999	
·21	·99822	·99985	·99997	1·00000	
·21	·99831	·99987	·99997		
·23	·99839	·99988	·99997		
·24	·99846	·99989	·99998		
2·25	0·99854	0·99989	0·99998		
·26	·99861	·99990	·99998		
·27	·99867	·99991	·99998		
·28	·99874	·99992	·99998		
·29	·99880	·99992	·99999		
2·30	0·99886	0·99993	0·99999		
·31	·99891	·99993	·99999		
·32	·99897	·99994	·99999		
·33	·99902	·99994	·99999		
·34	·99906	·99995	·99999		
2·35	0·99911	0·99995	0·99999		
·36	·99916	·99996	·99999		
·37	·99920	·99996	·99999		
·38	·99924	·99996	·99999		
·39	·99928	·99997	·99999		
2·40	0·99931	0·99997	1·00000		
·41	·99935	·99997			
·42	·99938	·99997			
·43	·99941	·99998			
·44	·99944	·99998			
2·45	0·99947	0·99998			
·46	·99950	·99998			
·47	·99952	·99998			
·48	·99955	·99998			
·49	·99957	·99999			
2·50	0·99959	0·99999			

m \ n	2	3
2·50	0·99959	0·99999
·51	·99961	·99999
·52	·99964	·99999
·53	·99965	·99999
·54	·99967	·99999
2·55	0·99969	0·99999
·56	·99971	·99999
·57	·99972	·99999
·58	·99974	·99999
·59	·99975	·99999
2·60	0·99976	0·99999
·61	·99978	·99999
·62	·99979	1·00000
·63	·99980	
·64	·99981	
2·65	0·99982	
·66	·99983	
·67	·99984	
·68	·99985	
·69	·99986	
2·70	0·99987	
·71	·99987	
·72	·99988	
·73	·99989	
·74	·99989	
2·75	0·99990	
·76	·99991	
·77	·99991	
·78	·99992	
·79	·99992	
2·80	0·99992	
·81	·99993	
·82	·99993	
·83	·99994	
·84	·99994	
2·85	0·99994	
·86	·99995	
·87	·99995	
·88	·99995	
·89	·99996	
2·90	0·99996	
·91	·99996	
·92	·99996	
·93	·99997	
·94	·99997	
2·95	0·99997	
·96	·99997	
·97	·99997	
·98	·99998	
·99	·99998	
3·00	0·99998*	

* 0·99999 is reached for $m = 3·07$; 1·00000 is reached for $m = 3·23$.

Table 2. *Percentage points of the probability integral of the mean deviation (m), with the population standard deviation as unit*

(*a*) Lower percentage points

Size of sample n	0·1 %	0·5 %	1·0 %	2·5 %	5·0 %	10·0 %
2	0·001	0·004	0·009	0·022	0·044	0·089
3	0·022	0·052	0·073	0·116	0·166	0·238
4	0·066	0·114	0·145	0·199	0·254	0·328
5	0·112	0·170	0·203	0·260	0·315	0·386
6	0·153	0·215	0·250	0·306	0·360	0·428
7	0·190	0·252	0·287	0·342	0·394	0·459
8	0·220	0·283	0·318	0·372	0·422	0·484
9	0·247	0·310	0·344	0·396	0·445	0·504
10	0·271	0·333	0·366	0·417	0·464	0·521
			Normal approximation:			
10	0·171	0·269	0·316	0·386	0·445	0·514

(*b*) Upper percentage points

Size of sample n	10·0 %	5·0 %	2·5 %	1·0 %	0·5 %	0·1 %
2	1·163	1·386	1·585	1·821	1·985	2·327
3	1·117	1·276	1·417	1·586	1·703	1·949
4	1·089	1·224	1·344	1·489	1·590	1·806
5	1·069	1·187	1·292	1·419	1·507	1·693
6	1·052	1·158	1·253	1·366	1·445	1·613
7	1·038	1·135	1·222	1·325	1·397	1·550
8	1·026	1·116	1·196	1·292	1·358	1·499
9	1·016	1·100	1·175	1·264	1·326	1·457
10	1·007	1·086	1·156	1·240	1·299	1·422
			Normal approximation:			
10	1·000	1·069	1·128	1·198	1·245	1·342

BOOK REVIEW

The Advanced Theory of Statistics. Vol. I. By MAURICE G. KENDALL. London: Charles Griffin and Co. Ltd., 1943. Pp. 457. Price 42*s*.

It is difficult to review the present volume without knowing precisely how the author will deal with the topics reserved for its promised successor. For although Mr Kendall expresses the hope that this first instalment can profitably be read before the publication of the second, it is clear, nevertheless, that the two parts will be complementary and that full justice can only be done to this first part after the two volumes have been considered together.

Mr Kendall defines his objective as the provision of a systematic treatment of statistical theory as it exists at the present time. The work is encyclopaedic and will receive little criticism on the grounds of what is omitted. It is not an elementary book, the various topics being all carried to an advanced stage and at times requiring of the reader considerable mathematical powers. As with advanced theoretical work in most of the sciences, the practical problems which originally suggested the discussions have often receded well into the background. This is not mentioned as a criticism, for the sole value of *scientific work does not necessarily lie in it being of immediate or even of ultimate practical importance.* It is, however, proper to point out that Mr Kendall is here, in the main, content to present us with a picture of statistical theory as he finds it. The more controversial job of assessing the value of the different parts of the structure, whether from a purely practical or from an aesthetic viewpoint, he leaves to the reader. Within these self-imposed limitations he has scored a notable success.

The first six chapters deal in some detail with the properties of frequency distributions. Chapter 3, entitled 'Moments and Cumulants', is here particularly satisfying. It develops concisely the general relationships between the various families of power statistics. The simple presentation of the transformations giving moments in terms of cumulants and vice versa will be welcomed, and the listing of a large number of the resulting formulae will enhance the value of the book as a reference work. In addition, the familiar corrections for grouping, due to Sheppard, are derived, and the rather subtle distinctions between the conditions necessary for their application and the conditions under which the so-called average corrections for grouping may be applied are clearly drawn.

Chapter 4, entitled 'Characteristic Functions', begins with a proof of the Inversion Theorem which states that the characteristic function uniquely determines the distribution function. It then discusses at length various theorems connected with the limits of infinite sequences of distribution functions, and with the so-called Problem of Moments, i.e. the problem of specifying conditions under which the moments determine uniquely the distribution function. As in discussions of the convergence of infinite series in pure mathematics, the interest in the limits of sequences of distribution functions is almost exclusively theoretical, and this chapter should appeal to anyone who enjoys himself in this type of work. The more practical question of determining approximations with the aid of only a few moments is considered in later chapters.

Chapter 5 introduces the simpler distributions which are of central importance in statistical practice. Chapter 6 continues with a description of the Pearson system of curves and of the series developments of the Normal and Poisson distributions, associated with the names of Gram, Charlier and Edgeworth. One welcomes the fact that both the Gram-Charlier and the Edgeworth developments from the normal distribution are given, for although these series may be but rearrangements of each other, it is the order and grouping of the terms that are all important when any practical applications are intended. Indeed, although this point is well made, it is of such importance that, even at the risk of appearing to labour it, the addition of some further numerical illustrations might have been useful.

The next five chapters deal with Theories of Probability, with Sampling and with Sampling Distributions. On the first topic the author rightly does not dogmatize, judging that the rest of the subject appears to go forward in much the same way, whatever are one's basic concepts as to the meaning of probability. This determination not to take sides in the present book is extended also to the problem of induction, where he minimizes the too violent contrasts which have in recent years been drawn between Bayes's Theorem and the Principle of Maximum Likelihood. Far from being diametrically opposed, Mr Kendall observes that, if some account is taken of the limiting processes by which continuous distributions are defined, and if Bayes's Postulate is introduced in an appropriate manner, the two principles have a very strong resemblance.

The chapter devoted to the so-called 'exact' sampling distributions follows familiar lines. Perhaps of greater interest is Chapter 11, which is entitled 'Approximations to Sampling Distributions'. Here

the problems associated with the distribution of the k-statistics are given a very full treatment. It is, moreover, an authoritative treatment since Mr Kendall has himself contributed so much to the elucidation of the methods which are here required. The k-statistics, which were introduced by R. A. Fisher in 1928, have the property that their expected values in repeated samples are the cumulants of the populations sampled. In general the exact distributions of the k-statistics cannot be derived, but the cumulants of the k-distributions can be obtained by following out certain rules which were given by Fisher. Mr Kendall describes these rules and gives full proofs of their validity. A large number of the formulae derived from applying the rules in particular cases are quoted for reference. This is a very live subject, and one feels that there is still some scope for development and simplification of these methods, for it must still be admitted that their successful application requires a high degree of virtuosity. Aesthetically, however, they are far in advance of the heavy algebraic manipulation demanded by the earlier approach. The mathematician will perhaps find most in the chapter to satisfy his sense of what has come to be called elegance in such studies.

After this the reviewer found the remainder of the book dealing with the χ^2 distribution and with the Theory of Correlation something of an anti-climax. Most of all at this stage one feels the need to refer to the promised second volume, to which has been assigned the general theory of regression analysis. Correlation and regression are so closely allied that even a temporary separation is distressing.

While the above remarks may give some idea of the nature of the subjects discussed in this first volume, they do not do justice to the thoroughness with which the author has accomplished his task. It is by no means a book to skim through. Mr Kendall notes rather ruefully in his preface that statistical theory is essentially mathematical and suggests that it is not easy to keep the mathematics from getting on top. He says, however, that he intends his work to be one on statistics and not on statistical mathematics. The distinction is a fine one, and I must confess to some difficulty in appreciating Mr Kendall's point here. Readers who are familiar with Dr Aitken's small work entitled *Statistical Mathematics* may perhaps ask whether his book ought not, for the same reasons, to be termed one on statistics and not on statistical mathematics. However, by whatever name we speak of this subject, anyone who can follow a mathematical argument, and who has also at least some small experience of practical statistical problems, will find plenty to reward him in a study of Mr Kendall's book.

B. L. WELCH

Vol. XXXIII. Part IV June, 1946

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY
BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

ASSISTED BY

EGON S. PEARSON

Reprinted by offset-litho

ISSUED BY THE BIOMETRIC LABORATORY
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

PRINTED IN GREAT BRITAIN

VOLUME XXXIII, PART IV JUNE 1946

STATISTICAL TECHNIQUES IN APPLIED PSYCHOLOGY

BY E. G. CHAMBERS, *From the Psychological Laboratory, Cambridge*

Psychological research work raises some statistical problems of a nature not usually encountered in biometric and economic studies, and the experimenter is frequently faced with serious difficulties in choosing suitable statistical methods for the treatment of his data. Naturally he wishes to make the treatment as exact and as fruitful as possible, and often the temptation to use modern methods, such as variate analysis or factor analysis, proves irresistible, notwithstanding the facts that the original data may be rather nebulous and that the results of statistical analysis have still to be interpreted in psychological terms. The question as to how far modern statistical techniques are legitimately applicable to psychological data is becoming increasingly important. This short paper is an attempt to indicate some of the difficulties involved and perhaps to interest statisticians in this field of endeavour.

The material collected by psychologists usually falls into one of a few categories. First there is the class of measurements. These are generally test scores, and though it may be begging an important psychological question to call them 'measurements' at all, yet little harm may be done by treating such data by the ordinary correlational and analytic techniques, provided always, of course, that the data satisfy the usual requirements of distribution, etc. Even here, however, the critical investigator will ask himself whether the more elaborate and imposing techniques really do add to the information gained from the use of simpler methods.

A second type of data consists of rankings. For example, a group of subjects may each be ranked according to his degree of possession of some psychological attribute or attributes. These rankings are usually made by one or more judges and are based on personal judgements. Now there is an increasing tendency for investigators to transform such ranked material into 'normally distributed' data, which are then subjected to product-moment correlation, variate analysis, or what you will. There are two commonly used methods of effecting this transformation. One method is to use the table giving scores for ordinal data in Fisher and Yates' *Statistical Tables for Biological, Agricultural and Medical Research*. In the other method the ranked data are divided into groups, the frequencies of which follow the normal scale more or less closely, and the groups are then allotted scores on a linear scale. For instance, in a recent piece of work the investigator divided a ranked group into seven subgroups containing respectively 5, 10, 20, 30, 20, 10 and 5 % of the individuals, and to these subgroups he then assigned the marks -3, -2, -1, 0, 1, 2 and 3. This then left him with a set of 'normally distributed' scores for some psychological attribute, which, with other similar sets, he used for producing a matrix of correlation coefficients, which in turn was subjected to factor analysis.

It seems to the present writer that here we have strayed a long way from the original ranked data, and that some of the steps taken are very difficult to justify. In the first place the original rankings were based on personal judgements, and we have no sort of guarantee that the judge was capable of making correct rankings for the psychological attributes under consideration or that he maintained a consistent standard of judgement throughout the

whole range. Assuming, however, that he was capable of making accurate assessments, we still cannot know what it actually was that he was assessing, even though he called it 'initiative' or 'conscientiousness' or whatever it was supposed to be. Further, in the absence of definite evidence from some other source, it is very doubtful whether we have the right to assume that psychological attributes are normally distributed in a selected population (the subjects in this instance were scholars at a particular school), so that the artificially produced set of 'normally distributed' scores may indeed have no counterpart in actuality. In view of these considerations it is extremely difficult to interpret in psychological terms any mathematical factors found in the matrix of correlation coefficients finally achieved.

There is another objection to normalizing ranked data which is not commonly realized. Unless the ranking is obtained by the use of some metric we cannot know that the intervals between successive ranks are equal; indeed, it is unlikely that they are. Errors of judgement will, however, tend to be equal at all points of the scale, so that the effect of normalizing the rankings will be to alter the relative numerical value of observational errors at different parts of the scale. Moreover, the variance of the normalized scores will not be the same as that of the original observational material, and in any analysis of this variance the effect we wish to isolate may have been distorted or even entirely masked by the process of normalizing.

A third type of psychological data is produced by getting some judge to assess individuals on a five- or ten-point scale according to their possession of some quality. It might be, for example, that a foreman in a factory is asked to assess his subordinates on their 'co-operativeness' or on their 'efficiency'. There are, of course, psychological difficulties involved in this process, but it is not the purpose of this paper to examine these. It is the way such data are treated statistically which is our concern here. Let us suppose that a group of workers are each assessed as A, B, C, D or E for 'efficiency', A signifying 'extremely efficient' and E 'extremely inefficient'. The question then frequently arises, how are these assessments related to the scores on some selective test? All too often this problem is tackled by transforming the literal grades into numerical scores by taking A as worth 5 marks, B as worth 4, and so on. These scores are then treated by any modern statistical technique that takes the investigator's fancy, frequently quite regardless of the fact that the numerical scores may be markedly leptokurtic or badly skewed in distribution. The nature of the true distribution of 'efficiency' and the fact that the assessments are the more or less imperfect judgements of someone who is usually untrained in making such judgements are points which are too often forgotten, the neatness of the mathematical techniques used lending a spurious appearance of accuracy to the whole proceeding.

The statistical treatment of these various sorts of psychological material is no mere academic matter but a vital practical problem, particularly at the present time when we are faced with rehabilitation and reorganization of labour on a large scale. Tests for industrial selection are becoming increasingly important, and it is essential to have some statistical methods of proving their validity. Unfortunately, it is extremely difficult to obtain adequate validating criteria from industry, and very often personal assessments of the sort described above are all that are available. This is a fact which cannot be burked, and in the writer's opinion no benefit is obtained by attempts to treat such assessments and rankings as other than what they are, particularly by attempts to transform them into exact numerical data in an artificially produced shape. There are, however, certain simple statistical methods which

do not make the assumptions involved in many modern techniques and whose use is not open to the objections briefly mentioned above. These methods are chiefly due to M. G. Kendall, sometimes in collaboration with B. Babington Smith, and have mostly been described in earlier issues of *Biometrika* (Kendall, 1938, 1942; Kendall & Babington Smith, 1939, 1940). They are a method of rank correlation, yielding a coefficient whose significance may readily be tested, the method of paired comparisons and a method of testing the agreement between several judges. These methods have already been used with fruitful results by the Unit for Applied Psychology at Cambridge in the field of industry. It is believed that a more reliable ranking of abilities and attributes may be obtained by the paired comparisons technique than by any other method, especially as the method carries with it its own estimate of a judge's consistency of judgement. Further, a psychologically untrained person may easily be able to compare pairs of individuals as regards some quality, whereas he would find it difficult if not impossible to rank all the members of even a relatively small group. The Kendall method of rank correlation has certain advantages over the Spearman method, since fresh material may be added from time to time without having to re-rank at each stage, and also since it allows the calculation of partial rank correlation coefficients.

An example of the use of these methods in dealing with an industrial problem may be of interest. A certain firm asked for help in the selection of foremen, preferably help in the form of a psychological test which the management itself could administer to candidates. The first stage in the inquiry was to seek information from those qualified to give an opinion as to the most important qualities involved in good foremanship. From the many suggestions made, six qualities were taken as being the most important requisites of good foremanship and the most representative of the general enlightened opinion. These qualities were:

(1) Ability to get on with the workers.
(2) Co-operation with the management.
(3) Technical knowledge.
(4) Organizing ability.
(5) Ability to maintain discipline.
(6) Initiative and improvisation.

The next step was to investigate how far existing foremen showed differences in respect of these qualities. Of the various possible ways of attempting this the method of allotting numerical scores for the degree of possession of each quality and the method of ranking the whole group of foremen for each quality were immediately rejected as unjustifiable and dangerous, since there was no way of checking the validity of such scores or rankings. The method of paired comparisons, however, seemed ideal for the purpose. There were ten foremen in the group and three judges were chosen who knew them all well enough to justify the making of comparisons between them. Each judge had to make 45 (i.e. $\frac{1}{2}n(n-1)$) comparisons between all possible pairs of foremen for each quality. The lists of comparisons were then examined for circular triads (e.g. A judged better than B, B better than C and C better than A), and coefficients of consistency calculated from the formula

$$\zeta = 1 - \frac{24d}{n^3 - 4n},$$

where d = number of triads (Kendall, 1943, p. 425). The results of this were as follows:

Quality	Judge *A*		Judge *B*		Judge *C*	
	d	ζ	d	ζ	d	ζ
1	0	1·0	0	1·0	3	0·925
2	0	1·0	2	0·951	3	0·925
3	0	1·0	0	1·0	1	0·975
4	1	0·975	6	0·851	1	0·975
5	0	1·0	0	1·0	6	0·851
6	1	0·975	2	0·951	4	0·911

This indicates that each judge was highly consistent in his judgements, especially judge *A*.

Next, the agreement between the three judges was examined by the calculation of a coefficient of agreement (Kendall, 1943, p. 427). The coefficient, u, is given by

$$u = \frac{2\Sigma}{\binom{m}{2}\binom{n}{2}} - 1,$$

where m = number of judges, n = number of objects judged, Σ = total number of agreements between judges. The significance of this coefficient is examined by calculating χ^2 and finding P for the appropriate number of degrees of freedom. P in this instance gives the probability that the observed value of Σ would be attained or exceeded by chance if preferences were assigned at random.

The results yielded were as under:

Quality	u	P
1	0·56	$<0·0001$
2	0·53	$<0·0001$
3	0·73	$<0·0001$
4	0·41	0·0001
5	0·20	0·014
6	0·64	$<0·0001$

On the whole, the three judges agreed with one another fairly well, except in the cases of quality 5 (Ability to maintain discipline), where the agreement is not good, and quality 4 (Organizing ability), where the agreement, though quite significant, is only fair.

A further method of comparing the agreement of the judges was possible. From the paired comparisons lists the ten foremen were ranked for each quality according to the judgements of each judge. These rankings were then correlated for each pair of judges, using Kendall's ranking method to produce τ coefficients. The following table shows the values of τ obtained, those in brackets being insignificant:

Quality	Judges *A* and *B*	Judges *A* and *C*	Judges *B* and *C*
1	0·60	0·52	0·57
2	0·57	0·48	0·67
3	0·69	0·68	0·83
4	(0·24)	0·50	0·46
5	(−0·07)	(0·20)	0·52
6	0·66	0·73	0·73

These coefficients confirm the findings of the previous table, showing that the agreement between the judges is good except in the cases of qualities 5 and 4.

In view of the reasonable agreement between the judges it was possible to obtain a combined ranking for each foreman for each quality by addition of the three ranks and re-ranking of the ten totals in each case. This is as far as this particular investigation, which is still in progress, has yet reached. The devising of a suitable psychological test for these qualities presents peculiar difficulties, and the test needs careful checking for reliability before assessing its value as a selective instrument. However, a reasonable criterion for various qualities needed in good foremanship is now available in this instance, and when the test rankings are finally obtained their association with the quality rankings may be examined.

By their nature, the statistical methods used in this example are applicable to small populations only. If some statistician could evolve modifications making them useful for larger groups or develop a method of combining results from several small groups, apart from averaging a number of values of τ, he would benefit the industrial psychologist enormously and help to rid psychological research of a very dangerous tendency to the indiscriminate use of elaborate analytical techniques. One other direction in which statistical research would be very welcome would be in the development of median statistics, for quite often in psychological work the nature and distribution of the data are such that means and standard deviations are almost meaningless.

REFERENCES

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81.

KENDALL, M. G. (1942). Partial rank correlation. *Biometrika*, **32**, 277.

KENDALL, M. G. (1943). *The Advanced Theory of Statistics*, **1**. London: Griffin and Co. Ltd.

KENDALL, M. G. (1945). *The Treatment of Ties in Ranking Problems. Biometrika*, **33**, 239.

KENDALL, M. G. & BABINGTON SMITH, B. (1939). The problem of m rankings. *Ann. Math. Statist.* **10**, 275.

KENDALL, M. G. & BABINGTON SMITH, B. (1940). On the method of paired comparisons. *Biometrika*, **31**, 324.

A USEFUL METHOD FOR THE ROUTINE ESTIMATION OF DISPERSION FROM LARGE SAMPLES

By A. E. JONES, *Rothamsted Experimental Station*

1. Introduction

It is often possible, in certain types of mass production, to use a large sample of articles for simple routine inspection and to find with ease the articles with more extreme values of the characteristic measured. Examples of this are articles which undergo a routine check on their length or weight, in which case extreme values can be sorted out either by sight, or by use of GO–NO GO checks on a balance set at two suitable weights.

In these cases, a great deal of labour can be saved, if the dispersion is estimated from these extreme values, which may comprise only about 5 % of the total. Such an estimate of dispersion may be used in controlling variability by specifying limits for this estimate. One method of specifying the variability, which avoids the complication of subdividing the sample, is to lay down limits for the difference between the sum of the r highest and r lowest values observed in the sample.

In this paper it will be shown how the mean, variance, and also higher moments of this difference can be found. Approximate formulae, which are reasonably easy to calculate, are given for the mean and variance of the difference. These should be satisfactory for most practical purposes. In Table 1 are given exact values of the mean and variance of the difference in the case when the parent population is Gaussian (normal) for selected sample sizes and values of r. The mean and variance with other parent populations may be calculated by applying equations (22) and (25) to Tables 3 and 4.

2. General formula for the mean

Let n independent observed values $x_1, x_2, \ldots, x_n$ form the sample and suppose $x_1, x_2, \ldots, x_n$ to be in decreasing order of magnitude.

Denote the r values greater than x_{r+1} by x'_i $(i = 1, \ldots, r)$ and the r values less than x_{n-r} by x''_j $(j = 1, \ldots, r)$. It should be noted that (x'_i) and (x''_j) are not themselves arranged in order of magnitude.

Assume the parent distribution to have a finite elementary probability law, say $f(x)$, for all x, such that the first two moments exist.

Let $p_1(x'_i|x_{r+1})$ and $p_2(x''_j|x_{n-r})$ be the elementary probability laws for x'_i and x''_j, given the values of x_{r+1} and x_{n-r}. Then

$$p_1(x'_i) = f(x'_i) \Big/ \int_{x_{r+1}}^{\infty} f(x)\,dx, \quad p_2(x''_j) = f(x''_j) \Big/ \int_{-\infty}^{x_{n-r}} f(x)\,dx. \tag{1}$$

Denoting the difference between the sums of the r highest and r lowest values by S,

$$S = \sum_{i=1}^{r} x'_i - \sum_{j=1}^{r} x''_j = S_1 - S_2,$$

where

$$S_1 = \sum_{i=1}^{r} x'_i, \quad S_2 = \sum_{j=1}^{r} x''_j.$$

From (1), $$E(x_i'|x_{r+1}) = \int_{x_{r+1}}^{\infty} xf(x)\,dx \Big/ \int_{x_{r+1}}^{\infty} f(x)\,dx = \frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}, \tag{2}$$

where $$\mu_k(x_{r+1}) = \int_{x_{r+1}}^{\infty} x^k f(x)\,dx.$$

Therefore $$E(x_i') = \int_{-\infty}^{\infty} \frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})} p(x_{r+1})\,dx_{r+1}, \tag{3}$$

where $p(x_{r+1})$ is the elementary probability law of x_{r+1}.

Now $$p(x_{r+1}) = \frac{n!}{r!(n-r-1)!}[\mu_0(x_{r+1})]^r[1-\mu_0(x_{r+1})]^{n-r-1} f(x_{r+1}). \tag{4}$$

Hence (3) becomes $$E(x_i') = \int_0^1 \frac{\mu_1(x_{r+1})}{\mu_0} \frac{n!}{r!(n-r-1)!} \mu_0^r(1-\mu_0)^{n-r-1}\,d\mu_0 \tag{5}$$

($\mu_1(x_{r+1})$ being expressed in terms of μ_0). Similarly

$$E(x_j'') = E\left[\frac{\nu_1(x_{n-r})}{\nu_0(x_{n-r})}\right] = \int_0^1 \frac{\nu_1(x_{n-r})}{\nu_0} \frac{n!}{r!(n-r-1)!} \nu_0^r(1-\nu_0)^{n-r-1}\,d\nu_0, \tag{6}$$

$\nu_1(x_{n-r})$ being supposed expressed in terms of ν_0, where

$$\nu_k(x_{n-r}) = \int_{-\infty}^{x_{n-r}} x^k f(x)\,dx.$$

Hence from (5) and (6), the expected value, or mean, of S is given by

$$\begin{aligned} E(S) = r\int_0^1 \frac{n!}{r!(n-r-1)!} \frac{\mu_1(x_{r+1})}{\mu_0} \mu_0^r(1-\mu_0)^{n-r-1}\,d\mu_0 \\ -r\int_0^1 \frac{n!}{r!(n-r-1)!} \frac{\nu_1(x_{n-r})}{\nu_0} \nu_0^r(1-\nu_0)^{n-r-1}\,d\nu_0. \end{aligned} \tag{7}$$

Thus, if the probability law $f(x)$ be known, the mean value of S can be obtained by numerical integration. In the special case of a symmetrical distribution with mean zero, we have

$$E(S) = 2r\int_0^1 \frac{n!}{r!(n-r-1)!} \frac{\mu_1(x_{r+1})}{\mu_0} \mu_0^r(1-\mu_0)^{n-r-1}\,d\mu_0.$$

Tables of $\mu_0(x)$, $\mu_1(x)$, $\mu_2(x)$ for a normally distributed variable are given in *Tables for Statisticians and Biometricians*, Pt. I, Table IX (K. Pearson, 1930). These considerably reduce the labour involved in computing $E(S)$ and have been used in the preparation of Table 1.

3. GENERAL FORMULA FOR VARIANCE

The variance of S may be obtained by a method similar to that used in finding the mean. Thus, first an expression is derived for the conditional value of the variance, x_{r+1} and x_{n-r} being fixed. The unconditioned variance is then obtained by taking the expected value of the conditional variance over variation of x_{r+1} and x_{n-r}. Thus

$$\begin{aligned} E[\{S-E(S)\}^2 \mid x_{r+1}, x_{n-r}] &= E[\{S_1 - E(S_1)\}^2 \mid x_{r+1}] \\ &+ E[\{S_2-E(S_2)\}^2 \mid x_{n-r}] - 2E[\{S_1-E(S_1)\}\{S_2-E(S_2)\} \mid x_{r+1}, x_{n-r}]. \end{aligned} \tag{8}$$

The first term on the right-hand side of (8) may be dealt with as follows:

$$\begin{aligned} &E[\{S_1-E(S_1)\}^2 \mid x_{r+1}] \\ &= E[\{S_1 - E(S_1 \mid x_{r+1})\}^2 \mid x_{r+1}] + \{E(S_1 \mid x_{r+1}) - E(S_1)\}^2 \\ &= E\left[\left\{\sum_{i=1}^{r}[x_i' - E(x_i' \mid x_{r+1})]\right\}^2 \middle| x_{r+1}\right] + r^2\{E(x_i' \mid x_{r+1}) - E(x_i')\}^2 \\ &= r \times [\text{Variance of } x_i',\ x_{r+1} \text{ being fixed}] + r^2\left\{\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})} - E\left[\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}\right]\right\}^2. \end{aligned} \tag{9}$$

Also, since $$E[\{x_i'-E(x_i'\,|\,x_{r+1})\}^2\,|\,x_{r+1}] = E(x_i'^2\,|\,x_{r+1})-[E(x_i'\,|\,x_{r+1})]^2$$

$$= \int_{x_{r+1}}^{\infty} x^2\left(\frac{f(x)}{\int_{x_{r+1}}^{\infty} f(x)\,dx}\right)dx - \frac{\mu_1^2(x_{r+1})}{\mu_0^2(x_{r+1})}$$

$$= \frac{\mu_2(x_{r+1})}{\mu_0(x_{r+1})} - \frac{\mu_1^2(x_{r+1})}{\mu_0^2(x_{r+1})}, \tag{10}$$

equation (9) becomes

$$E[\{S_1-E(S_1)\}^2\,|\,x_{r+1}] = r\left\{\frac{\mu_2(x_{r+1})}{\mu_0(x_{r+1})}-\frac{\mu_1^2(x_{r+1})}{\mu_0^2(x_{r+1})}\right\} + r^2\left\{\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})} - E\left[\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}\right]\right\}^2 \tag{11}$$

Similarly, we obtain for the second term on the right-hand side of (8):

$$E[\{S_2-E(S_2)\}^2\,|\,x_{n-r}] = r\left\{\frac{\nu_2(x_{n-r})}{\nu_0(x_{n-r})}-\frac{\nu_1^2(x_{n-r})}{\nu_0^2(x_{n-r})}\right\} + r^2\left\{\frac{\nu_1(x_{n-r})}{\nu_0(x_{n-r})} - E\left[\frac{\nu_1(x_{n-r})}{\nu_0(x_{n-r})}\right]\right\}^2 \tag{12}$$

If x_{r+1} and x_{n-r} are fixed, S_1 and S_2 are independent and so

$$E[\{S_1-E(S_1)\}\{S_2-E(S_2)\}\,|\,x_{r+1},x_{n-r}] = \{E[S_1\,|\,x_{r+1}]-E(S_1)\}\{E[S_2\,|\,x_{n-r}]-E(S_2)\}. \tag{13}$$

Substituting (11), (12) and (13) in (8) and integrating over the joint probability distribution of x_{r+1} and x_{n-r}, we obtain the unconditioned expected value

$$E[\{S-E(S)]^2\} = E\big[E[\{S-E(S)\}^2\,|\,x_{r+1},x_{n-r}]\big]$$

$$= E\left[r\left(\frac{\mu_2}{\mu_0}-\frac{\mu_1^2}{\mu_0^2}\right)+r^2\left\{\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right)\right\}^2\right] + E\left[r\left(\frac{\nu_2}{\nu_0}-\frac{\nu_1^2}{\nu_0^2}\right)+r^2\left\{\frac{\nu_1}{\nu_0}-E\left(\frac{\nu_1}{\nu_0}\right)\right\}^2\right]$$

$$-2r^2E\left[\left\{\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right)\right\}\left\{\frac{\nu_1}{\nu_0}-E\left(\frac{\nu_1}{\nu_0}\right)\right\}\right], \tag{14}$$

where μ_0, μ_1, μ_2, ν_0, ν_1, ν_2 have been written in place of $\mu_0(x_{r+1})$, $\mu_1(x_{r+1})$, $\mu_2(x_{r+1})$, $\nu_0(x_{n-r})$, $\nu_1(x_{n-r})$, $\nu_2(x_{n-r})$ for conciseness.

The first term on the right-hand side of (14) is the expectation of a function of x_{r+1} only, and can therefore be expressed as a single integral. Similarly, the second term is a function of x_{n-r} only and can also be expressed as a single integral.

The third term which arises from the correlation between x_{r+1} and x_{n-r} is a double integral. However, its value can be estimated roughly by the following method which also indicates that the absolute magnitude of this term is small provided r/n is small.

It is known that, using the same abbreviations as in (14), the joint elementary probability law of $\mu_0(x_{r+1})$ and $\nu_0(x_{n-r})$ is

$$p(\mu_0,\nu_0) = \frac{n!}{(r!)^2\,(n-2r-2)!}\,\mu_0^r\nu_0^r(1-\mu_0-\nu_0)^{n-2r-2}. \tag{15}$$

By taking logarithms and expanding μ_0 and ν_0 about their respective means we obtain the approximation

$$p(\mu_0,\nu_0) \simeq \exp\left[-\frac{(n-2)^2}{2}\left\{(u_1^2+u_2^2)\left(\frac{1}{r}+\frac{1}{n-2r-2}\right)+\frac{2u_1u_2}{n-2r-2}\right\}\right],$$

where $$u_1 = \mu_0-E(\mu_0) \quad\text{and}\quad u_2 = \nu_0-E(\nu_0).$$

Hence the correlation coefficient of u_1 and u_2 (or μ_0 and ν_0) is very nearly $-r/(n-r)$. Also to the first order

$$\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right) \propto -u_1, \quad \frac{\nu_1}{\nu_0}-E\left(\frac{\nu_1}{\nu_0}\right) \propto +u_2.$$

Therefore

$$E\left[\left\{\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right)\right\}\left\{\frac{\nu_1}{\nu_0}-E\left(\frac{\nu_1}{\nu_0}\right)\right\}\right] \simeq \frac{r}{n-r}\sqrt{\left[\left(\text{Variance of }\frac{\mu_1}{\mu_0}\right)\times\left(\text{Variance of }\frac{\nu_1}{\nu_0}\right)\right]}. \tag{16}$$

Evidently the effect of the correlation between x_{r+1} and x_{n-r} will be small if r/n is small. If the approximate expression (16) is used in (14) the resulting accuracy should be adequate.

From (14) and (16) we have

$$\text{Variance of } S \simeq r(G_1+G_2)+r(r-1)(H_1+H_2)-\frac{2r^3}{n-r}\sqrt{(H_1H_2)}, \tag{17}$$

where

$$G_1=\int_0^1\left[\frac{\mu_2(x_{r+1})}{\mu_0}-\left\{E\left[\frac{\mu_1(x_{r+1})}{\mu_0}\right]\right\}^2\right]\frac{n!}{r!(n-r-1)!}\mu_0^r(1-\mu_0)^{n-r-1}\,d\mu_0, \tag{18·1}$$

$$G_2=\int_0^1\left[\frac{\nu_2(x_{n-r})}{\nu_0}-\left\{E\left[\frac{\nu_1(x_{n-r})}{\nu_0}\right]\right\}^2\right]\frac{n!}{r!(n-r-1)!}\nu_0^r(1-\nu_0)^{n-r-1}\,d\nu_0, \tag{18·2}$$

$$H_1=\int_0^1\left\{\frac{\mu_1(x_{r+1})}{\mu_0}-E\left[\frac{\mu_1(x_{r+1})}{\mu_0}\right]\right\}^2\frac{n!}{r!(n-r-1)!}\mu_0^r(1-\mu_0)^{n-r-1}\,d\mu_0, \tag{18·3}$$

$$H_2=\int_0^1\left\{\frac{\nu_1(x_{n-r})}{\nu_0}-E\left[\frac{\nu_1(x_{n-r})}{\nu_0}\right]\right\}^2\frac{n!}{r!(n-r-1)!}\nu_0^r(1-\nu_0)^{n-r-1}\,d\nu_0. \tag{18·4}$$

Of course, as in equations (5) and (6), functions of x_{r+1}, x_{n-r} in the above four expressions [e.g. $\mu_2(x_{r+1})$, $\nu_2(x_{n-r})$] are supposed to be expressed in terms of μ_0 and ν_0 [i.e. $\mu_0(x_{r+1})$ and $\nu_0(x_{n-r})$].

The higher order moments may also be evaluated by the same method. The work involved unfortunately becomes heavy for practical purposes. The third moment of S_1 about its mean, for example, is

$$E[\{S_1-E(S_1)\}^3]$$
$$=rE\left[\frac{\mu_3}{\mu_0}-3\frac{\mu_2}{\mu_0}\frac{\mu_1}{\mu_0}+2\frac{\mu_1^3}{\mu_0^3}\right]+r^2E\left[\left\{\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right)\right\}\left\{\frac{\mu_2}{\mu_0}-\frac{\mu_1^2}{\mu_0^2}\right\}\right]+r^3E\left[\left\{\frac{\mu_1}{\mu_0}-E\left(\frac{\mu_1}{\mu_0}\right)\right\}^3\right]. \tag{19}$$

The Moment Generating Function can be written

$$E(e^{S_1t})=\int_{-\infty}^{\infty}\left[\int_{x_{r+1}}^{\infty}e^{xt}f(x)\,dx\right]^r\left[1-\int_{x_{r+1}}^{\infty}f(x)\,dx\right]^{n-r-1}f(x_{r+1})\,dx_{r+1}.$$

It seems likely that for population distributions likely to occur in practice, the distribution of S will tend to the normal law as r becomes large, provided r/n remains small.

4. Further approximations for mean and variance of S

First consider the well-known equality

$$\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\,dy=\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Differentiating both sides with respect to α we have

$$\int_0^1 \log y\, y^{\alpha-1}(1-y)^{\beta-1}\,dy=-\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}\sum_{k=\alpha}^{\alpha+\beta-1}\frac{1}{k},$$

provided α and β are integers. Differentiating once again

$$\int_0^1 (\log y)^2 y^{\alpha-1}(1-y)^{\beta-1}\,dy=\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}\left\{\left(\sum_{k=\alpha}^{\alpha+\beta-1}\frac{1}{k}\right)^2+\sum_{k=\alpha}^{\alpha+\beta-1}\frac{1}{k^2}\right\}.$$

Hence $$\int_0^1 \left(\log y + \sum_{k=\alpha}^{\alpha+\beta-1} \frac{1}{k}\right)^2 y^{\alpha-1}(1-y)^{\beta-1}\,dy = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} \sum_{k=\alpha}^{\alpha+\beta-1} \frac{1}{k^2},$$

and in general

$$\int_0^1 \left(\log y + \sum_{k=\alpha}^{\alpha+\beta-1} \frac{1}{k}\right)^t y^{\alpha-1}(1-y)^{\beta-1}\,dy = (-1)^t\,(t-1)!\,\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} \sum_{k=\alpha}^{\alpha+\beta-1} \frac{1}{k^t}.$$

Now consider the variable $$L = \log \mu_0(x_{r+1}).$$

Since $$p(\mu_0) = \frac{n!}{r!\,(n-r-1)!}\mu_0^r(1-\mu_0)^{n-r-1},$$

it follows from the results just obtained that the central moments of L are

$$\bar{L} = E(L) = -\sum_{k=r+1}^{n} \frac{1}{k}, \tag{20·1}$$

$$E[(L-\bar{L})^2] = \sum_{k=r+1}^{n} \frac{1}{k^2}, \tag{20·2}$$

$$E[(L-\bar{L})^t] = (-1)^t\,(t-1)!\sum_{k=r+1}^{n} k^{-t}. \tag{20·3}$$

The approximations which we shall now obtain for the mean and variance of S apply to cases where the population distribution can be represented approximately as descending exponentially in the region of $\mu_0 = r/n$ for x increasing and in the region of $\nu_0 = r/n$ for x decreasing. They will be obtained by expanding the functions $\dfrac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}$, $\dfrac{\mu_2(x_{r+1})}{\mu_0(x_{r+1})}$, etc., as Taylor series in $L = \log \mu_0(x_{r+1})$ about the expected value ($\bar{L}$) of L.

Defining ξ by the equation $\bar{L} = \log \mu_0(\xi)$ we have, neglecting second and higher order terms,

$$\frac{\mu_1}{\mu_0} \simeq \left(\frac{\mu_1}{\mu_0}\right)_{x_{r+1}=\xi} + \left[-\left(\frac{\mu_1}{\mu_0}\right) + \xi\right](L-\bar{L}). \tag{21}$$

By (3) $$E(x_i') = E\left[\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}\right].$$

Hence from (21) $$E(x_i') \simeq \left(\frac{\mu_1}{\mu_0}\right)_{x_{r+1}=\xi},$$

where $$\log \mu_0(\xi) = \bar{L} = -\sum_{k=r+1}^{n} \frac{1}{k},$$

i.e. $$\mu_0(\xi) = \int_\xi^\infty f(x)\,dx = \exp\left(-\sum_{k=r+1}^{n} \frac{1}{k}\right). \tag{22}$$

Similarly, $$E(x_\gamma'') = E\left[\frac{\nu_1(x_{n-r})}{\nu_0(x_{n-r})}\right] \simeq \left(\frac{\nu_1}{\nu_0}\right)_{x_{n-r}=\eta},$$

where $$\nu_0(\eta) = \int_{-\infty}^{\eta} f(x)\,dx = \exp\left(-\sum_{k=r+1}^{n} \frac{1}{k}\right). \tag{23}$$

So $$E(S) \simeq r\left[\left(\frac{\mu_1}{\mu_0}\right)_{x_{r+1}=\xi} - \left(\frac{\nu_1}{\nu_0}\right)_{x_{n-r}=\eta}\right]. \tag{24}$$

This provides an approximate formula for the mean of S.

To obtain an approximate formula for the variance of S first consider the terms on the right-hand side of equation (17). Neglecting terms of higher order than the second, and remembering that $E(L-\bar{L}) = 0$, we have

$$G_1 = E\left[\frac{\mu_2(x_{r+1})}{\mu_0(x_{r+1})}\right] - \left\{E\left[\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}\right]\right\}^2 \simeq \left(\frac{\mu_2}{\mu_0} - \frac{\mu_1^2}{\mu_0^2}\right)_{x_{r+1}=\xi} + \frac{1}{2}\left[\frac{\mu_2}{\mu_0} - \frac{\mu_1^2}{\mu_0^2} - \left(\frac{\mu_1}{\mu_0} - \xi\right)^2 + \frac{2(\mu_1 - \xi\mu_0)}{f(\xi)}\right]_{x_{r+1}=\xi} E[(L-\bar{L})^2],$$

i.e.
$$G_1 \simeq \left(\frac{\mu_2}{\mu_0} - \frac{\mu_1^2}{\mu_0^2}\right)_{x_{r+1}=\xi} + \frac{1}{2}\left(\sum_{k=r+1}^{n}\frac{1}{k^2}\right)\left[\frac{\mu_2}{\mu_0} - \frac{\mu_1^2}{\mu_0^2} - \left(\frac{\mu_1}{\mu_0} - \xi\right)^2 + \frac{2(\mu_1 - \xi\mu_0)}{f(\xi)}\right]_{x_{r+1}=\xi}.$$

Also from (18) and (19)

$$H_1 = E\left[\left\{\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})} - E\left[\frac{\mu_1(x_{r+1})}{\mu_0(x_{r+1})}\right]\right\}^2\right] \simeq \left(\frac{\mu_1}{\mu_0} - \xi\right)^2_{x_{r+1}=\xi}\sum_{k=r+1}^{n}\frac{1}{k^2}.$$

Hence
$$rG_1 + r(r-1)H_1 \simeq r\left(\frac{\mu_2}{\mu_0} - \frac{\mu_1^2}{\mu_0^2}\right)_{x_{r+1}=\xi}\left(1 + \frac{1}{2}\sum_{k=r+1}^{n}\frac{1}{k^2}\right) + r(r-\tfrac{3}{2})\left[\frac{\mu_1(\xi)}{\mu_0(\xi)} - \xi\right]^2\sum_{k=r+1}^{n}\frac{1}{k^2} + \frac{r[\mu_1(\xi) - \xi\mu_0(\xi)]}{f(\xi)}\sum_{k=r+1}^{n}\frac{1}{k^2}. \quad (25)$$

The last term on the right-hand side of (25) involves $f(\xi)$ which will be generally rather difficult to estimate, since ξ will be in the tail of the distribution. It will therefore be desirable to be able to make some approximation to this term. The following approximation, which may be useful, is based on the assumption of an exponential rate of decrease of the probability density as x increases from ξ:

$$\frac{\mu_1(\xi) - \xi\mu_0(\xi)}{f(\xi)} = \left[\frac{\mu_1(\xi)}{\mu_0(\xi)} - \xi\right]\frac{\int_\xi^\infty f(x)\,dx}{f(\xi)} = \left[\frac{\mu_1(\xi)}{\mu_0(\xi)} - \xi\right]\frac{\int_\xi^\infty (x-\xi)f(x)\,dx}{\int_\xi^\infty f(x)\,dx} \simeq \left[\frac{\mu_1(\xi)}{\mu_0(\xi)} - \xi\right]^2.$$

From (23) we have

$$rG_1 + r(r-1)H_1 \simeq r\left[\frac{\mu_2(\xi)}{\mu_0(\xi)} - \frac{\mu_1^2(\xi)}{\mu_0^2(\xi)}\right]\left(1 + \frac{1}{2}\sum_{k=r+1}^{n}\frac{1}{k^2}\right) + r(r-\tfrac{1}{2})\left[\frac{\mu_1(\xi)}{\mu_0(\xi)} - \xi\right]^2\sum_{k=r+1}^{n}\frac{1}{k^2}. \quad (26)$$

Similarly, approximate expressions may be obtained for G_2 and H_2. So from (17),

$$\text{Variance of } S = r\left[\frac{\mu_2(\xi)}{\mu_0(\xi)} - \frac{\mu_1^2(\xi)}{\mu_0^2(\xi)} + \frac{\nu_2(\eta)}{\nu_0(\eta)} - \frac{\nu_1^2(\eta)}{\nu_0^2(\eta)}\right]\left(1 + \frac{1}{2}\sum_{k=r+1}^{n}\frac{1}{k^2}\right) + r\sum_{k=r+1}^{n}\frac{1}{k^2}\left[(r-\tfrac{1}{2})\left(\frac{\mu_1[\xi]}{\mu_0[\xi]} - \xi\right)^2 + (r-\tfrac{1}{2})\left(\frac{\nu_1[\eta]}{\nu_0[\eta]} - \eta\right)^2 + \frac{2r^2}{n-r}\left(\frac{\mu_1[\xi]}{\mu_0[\xi]} - \xi\right)\left(\frac{\nu_1[\eta]}{\nu_0[\eta]} - \eta\right)\right]. \quad (27)$$

It will be seen that only a knowledge of the tails of the parent distribution from $-\infty$ to η, and ξ to $+\infty$ is required to evaluate equations (24) and (27). Also, $(\mu_0(\xi) - r/n)$ and $(\nu_0(\eta) - r/n)$ will be positive and fairly small.

5. Practical application

It would be inadvisable to use this method, except in cases when approximations (24) and (27) apply, i.e. when the parent probability density function decreases steadily in each tail of the distribution. Also, it would be rather difficult, as equation (17) requires considerable computation.

Table 1. *Mean and standard error of S for a normal distribution with unit standard deviation*

r \ n	100	200	400	600	800	1000
5	**20·2*** 1·69*	**23·0*** 1·56*	**25·6** 1·45	**26·9*** 1·39*	**27·9** 1·36	**28·6*** 1·33*
7	**26·5** 2·09	**30·5** 1·92	**34·2** 1·78	**36·2** 1·71	**37·6** 1·67	**38·7** 1·63
10	—	**40·9*** 2·41*	**46·4** 2·22	**49·4** 2·13	**51·4** 2·07	**53·0*** 2·03*
12	—	**47·3** 2·70	**54·1** 2·49	**57·7** 2·38	**60·2** 2·31	**62·1** 2·26
16	—	—	**68·5** 2·98	**73·6** 2·84	**77·1** 2·76	**79·7** 2·70
20	—	—	**82·1** 3·42	**88·7** 3·26	**93·1** 3·16	**96·5*** 3·09*

The mean is written in bold type and the standard error in normal type below it.

* This indicates those figures which have been checked by exact computations.

Table 2. $\dfrac{\textit{Standard error of } S}{\textit{Mean of } S} \times 100$ *for samples from a normal distribution*

r \ n	100	200	400	600	800	1000
5	8·4	6·8	5·7	5·2	4·9	4·7
7	7·9	6·3	5·2	4·7	4·4	4·2
10	—	5·9	4·8	4·3	4·0	3·8
12	—	5·7	4·6	4·1	3·8	3·6
16	—	—	4·4	3·9	3·6	3·4
20	—	—	4·2	3·7	3·4	3·2
Corresponding ratio for s*	7·1	5·0	3·5	2·9	2·5	2·2

* The bottom row gives $100 \times$ standard error/mean, for the standard derivation s when calculated as $s = \sqrt{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)\right]}$ from a random, normal sample of size n.

Table 1 gives values of the mean and standard error of S for a normal distribution with unit variance. It will be seen that efficiency* is not much improved by increasing r/n beyond 4 %. Table 1 has been mostly evaluated from the formulae (24) and (27), but a number of exact computations have been made and these show that the approximations are accurate to a unit in the last figure shown in Table 1. More precisely, the maximum error was 0·2 % in the mean and 0·7 % in the standard deviation of S.

* Efficiency: If E is the efficiency of an estimate of the standard deviation made from a sample of size, n, then the best possible estimate of the standard deviation from a sample of size, nE, would have the same accuracy as measured by its standard error.

Table 3. *Values of* $\exp\left(-\sum_{k=r+1}^{n}\frac{1}{k}\right) = \mu_0(\xi)$

r \ n	100	200	400	600	800	1000
5	0·05480	0·02741	0·01376	0·00918	0·00689	0·00551
6	0·06474	0·03245	0·01626	0·01084	0·00813	0·00651
7	0·07468	0·03743	0·01876	0·01251	0·00938	0·00751
8	0·08463	0·04242	0·02125	0·01417	0·01063	0·00851
9	—	0·04740	0·02375	0·01584	0·01188	0·00951
10	—	0·05239	0·02625	0·01751	0·01313	0·01051
12	—	0·06236	0·03124	0·02084	0·01563	0·01251
14	—	0·07233	0·03624	0·02417	0·01813	0·01451
16	—	—	0·04124	0·02753	0·02063	0·01651
18	—	—	0·04623	0·03084	0·02313	0·01851
20	—	—	0·05123	0·03417	0·02563	0·02051

Table 4. *Values of* $\sum_{k=r+1}^{n}\frac{1}{k^2}$

r \ n	100	200	400	600	800	1000
5	0·1714	0·1763	0·1788	0·1797	0·1801	0·1803
6	0·1436	0·1486	0·1510	0·1519	0·1523	0·1526
7	0·1232	0·1282	0·1306	0·1315	0·1319	0·1321
8	0·1076	0·1125	0·1150	0·1158	0·1163	0·1165
9	—	0·1002	0·1027	0·1035	0·1039	0·1042
10	—	0·0902	0·0927	0·0935	0·0939	0·0942
12	—	0·0750	0·0775	0·0783	0·0787	0·0790
14	—	0·0640	0·0664	0·0673	0·0677	0·0679
16	—	—	0·0581	0·0589	0·0593	0·0596
18	—	—	0·0515	0·0524	0·0529	0·0532
20	—	—	0·0463	0·0471	0·0475	0·0478

When sampling from a normal population (in so far as the values of n and r tabled are appropriate), an estimate of the standard deviation, σ, can be obtained by calculating S from the sample and dividing it by the mean S given in Table 1. The standard error of this estimate, expressed as a percentage of the true σ, is given in Table 2. The percentage error of the usual estimate of σ based on the sums of squares of the n observations, which is approximately equal to $100/\sqrt{(2n)}$, is shown at the bottom of the table.

The exact parent probability distribution is, however, usually unknown and in order to estimate the mean and variance of S (*the difference between the r highest and r lowest values*) in a sample of n, a grand sample at least ten times as large, is required. If this grand sample, say of m values, is available, the mean and variance of S in samples of n may then be estimated as follows:

(a) Given n and r, Table 3 shows the corresponding values of $\mu_0(\xi)$ and $\nu_0(\eta)$, the quantities defined in equations (22) and (23), for a number of values of n.

(*b*) As the parent distribution is unknown, ξ and η cannot be found from $\mu_0(\xi)$ and $\nu_0(\eta)$. However, out of the grand sample about $p' = m\mu_0$ observations may be expected to be greater than ξ and about p' less than η.

(*c*) Denote the nearest integer to p' by p. From the grand sample, find the p largest values —call them y'_i $(i = 1, 2, \ldots, p)$—and the p smallest values, y''_j $(j = 1, 2, \ldots, p)$. Denote the $(p+1)$th value (from the highest) by ξ and the $(n-p)$th by η.

Calculate the mean and variance of the set of values y'_i $(i = 1, 2, \ldots, p)$. Call these M' and V'. Similarly, find the mean and variance of y''_j, and denote these by M'' and V''. Then, using equations (24) and (27),

$$\text{Mean of } S = r(M' - M''). \tag{28}$$

$$\text{Variance of } S = r(V' + V'')\left(1 + \frac{1}{2}\sum_{k=r+1}^{n}\frac{1}{k^2}\right) + r\sum_{k=r+1}^{n}\frac{1}{k^2}\left[(r-\tfrac{1}{2})(M'-\xi)^2 + (r-\tfrac{1}{2})(\eta - M'')^2 - \frac{2r^2}{n-r}(M'-\xi)(\eta - M'')\right]. \tag{29}$$

Values of $\sum\limits_{k=r+1}^{n}\frac{1}{k^2}$ are given in Table 4.

I would like to thank Mr N. L. Johnson for his help in preparing this paper for publication.

INEQUALITIES IN TERMS OF MEAN RANGE

By C. B. WINSTEN

CONTENTS

I. An inequality holding for any frequency distribution

Usually in statistical work the form of a frequency distribution is known, or assumed, so that it is possible to calculate exactly the fraction of the distribution lying in a given interval. It may happen, however, that though the mean, μ, and standard deviation, σ, have been estimated from sufficient data for errors of sampling to be neglected, nothing else is known about the distribution. Even in this case we know by Tchebycheff's inequality that the interval $(\mu - t\sigma, \mu + t\sigma)$ does not contain less than a fraction $1 - 1/t^2$ of the distribution, for any $t > 1$.

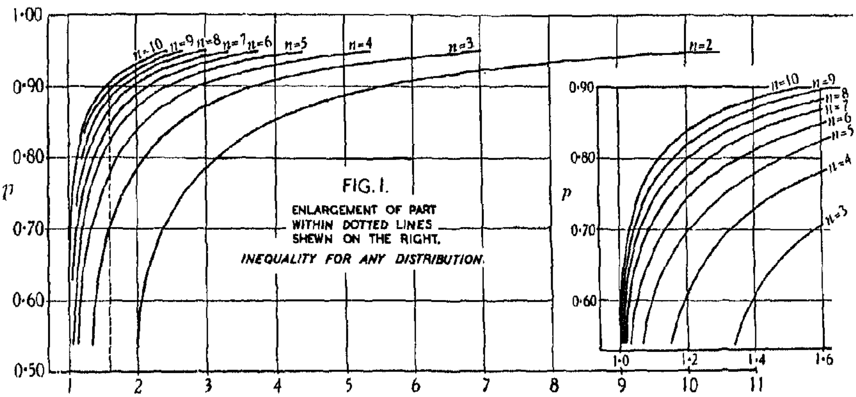


FIG. 1.
ENLARGEMENT OF PART WITHIN DOTTED LINES SHEWN ON THE RIGHT.
INEQUALITY FOR ANY DISTRIBUTION.

If, instead of the standard deviation of the distribution, the mean range of samples of n, w_n is known (such a situation is likely, for example, with standard control chart procedure), a different inequality is required, and this is given below. The new inequality is not an exact analogue of Tchebycheff's. Instead of considering the fixed interval $(\mu - t\sigma, \mu + t\sigma)$, we consider a variable interval $(x, x + tw_n)$ of fixed length $L = tw_n$. For a given distribution d, and fixed t, the fraction of d falling inside this interval is a function of x, $p_d(x, t)$ say. We know that $p_d(x, t) \leqslant 1$, and it therefore follows that $p_d(x, t)$ has, for all x, an upper bound, $p_d(t)$ say. It is this upper bound, $p_d(t)$, that will be considered.

The actual expression of the inequality is not so simple as Tchebycheff's. In practice, therefore, it is easier to use Fig. 1, as follows. Suppose the length of the interval, L, is given, and also the mean range for samples of n, w_n ($n = 2, 3, \ldots, 9, 10$). First find $t = L/w_n$. Then

find the ordinate of the appropriate curve in Fig. 1 with this value of t as abscissa. Suppose this ordinate is $p(t)$.

Then for any frequency distribution whatsoever, $p_d(t) > p(t)$.

As an example, suppose we have found that $w_5 = 3{\cdot}20$ cm. and $L = 6{\cdot}57$ cm. Then $t = 6{\cdot}57/3{\cdot}20 = 2{\cdot}05$, so that, from Fig. 1, $p(t) = 0{\cdot}88$. Thus, for practical purposes, we know that it is possible to choose an interval of length 6·57 cm. so that it will include at least 88 % of the distribution, but the inequality does not tell us *how* to choose this interval.

It may be desirable to have a more accurate estimate of $p(t)$ than can be obtained from the figure. In that case Table 1 can be used. In that table, for some values of a variable y, the values of a function $1/R_n(y)$ are given. (For the definition of $R_n(y)$ see § III.) Now $1/R_n(y) = t$, and $y = 1-p$ for the inequality we are considering. Hence we proceed as follows: for the given value of $t = 1/R_n(y)$ find the corresponding value of y by an inverse interpolation from the table; then find the value of $p(t)$ by subtracting the value of y obtained from unity. It is more accurate to use harmonic inverse interpolation to obtain the value of y.

Table 1. $\dfrac{1}{R_n(y)}$

$$\frac{1}{R_n(y)}\begin{cases} = t \text{ in general inequality,} \\ = 2t \text{ in inequality for unimodal symmetrical distributions}\end{cases}$$

y \ n	2	3	4	5	6	7	8	9	10
0·45	2·020	1·347	1·153	1·074	1·037	1·019	1·010	1·005	1·003
0·40	2·083	1·389	1·184	1·097	1·054	1·031	1·018	1·010	1·006
0·35	2·198	1·465	1·240	1·138	1·084	1·052	1·033	1·021	1·013
0·30	2·381	1·587	1·330	1·206	1·134	1·090	1·061	1·042	1·029
0·25	2·667	1·778	1·471	1·313	1·217	1·154	1·111	1·081	1·059
0·225	2·867	1·912	1·571	1·389	1·277	1·202	1·150	1·112	1·085
0·20	3·125	2·083	1·698	1·488	1·355	1·265	1·202	1·155	1·120
0·175	3·462	2·309	1·866	1·619	1·461	1·352	1·273	1·215	1·171
0·15	3·922	2·614	2·094	1·798	1·606	1·472	1·375	1·301	1·245
0·125	4·571	3·048	2·418	2·053	1·814	1·647	1·523	1·430	1·357
0·10	5·556	3·704	2·909	2·442	2·134	1·917	1·755	1·632	1·535
0·075	7·207	4·805	3·733	3·098	2·677	2·378	2·155	1·983	1·847
0·05	10·526	7·018	5·391	4·421	3·775	3·315	2·971	2·705	2·492

For example, suppose $n = 6$, $t = 1{\cdot}658$. The two nearest values of $t = 1/R_n(y)$ in Table 1, and the corresponding values of y and $1/y$ are:

$$t_1 = 1{\cdot}606, \qquad y_1 = 0{\cdot}150, \qquad \frac{1}{y_1} = 6{\cdot}667,$$

$$t_2 = 1{\cdot}814, \qquad y_2 = 0{\cdot}125, \qquad \frac{1}{y_2} = 8{\cdot}000.$$

Since $\dfrac{t-t_1}{t_2-t_1} = \dfrac{1}{4}$ and $\dfrac{1}{y_2} - \dfrac{1}{y_1} = \dfrac{4}{3}$, we find that $\dfrac{1}{y} = 7{\cdot}000$, so that $y = 0{\cdot}143$ and the value of p required is 0·857.

The inequality given here is the best possible of its type. This implies that, if all that is known about a distribution is the mean range for samples of n, for one and only one value of n, then for each t there is some distribution giving values of $p_d(t)$ as near to the value of $p(t)$

given by Fig. 1 as we please. Of course, if we know something more about a distribution, e.g. if we know both w_4 and w_5, it might be possible to find closer limits to $p_d(t)$.

The equations of the curves drawn in Fig. 1, and their derivation, are given in § III. For values of t larger than those given in Fig. 1, the approximate formula $p(t) = 1 - 1/(nt)$ is fairly satisfactory, provided n is small. For values of n greater than 10, $p(t)$ can be found from the equation given in § III.

II. An inequality holding for unimodal symmetrical distributions

If the mean, μ, and standard deviation, σ, of a distribution are known, and, in addition, the distribution is known to be unimodal and symmetrical, an inequality can be used which is a considerable improvement on Tchebycheff's. The simplest form of the Gauss-Winkler inequality states that, for a unimodal symmetrical distribution, the interval $(\mu - t\sigma, \mu + t\sigma)$ will contain not less than $1 - 4/(9t^2)$ of the distribution, for sufficiently large t. In fact there are

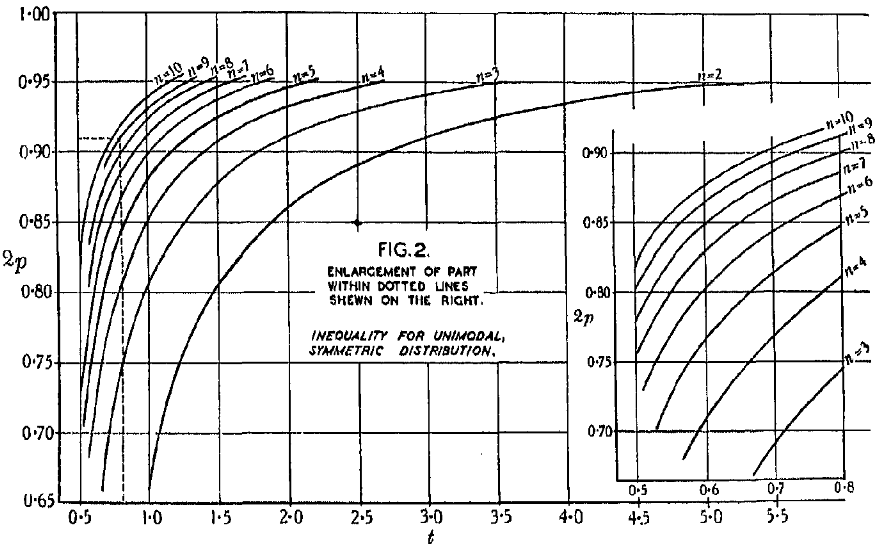


FIG. 2. ENLARGEMENT OF PART WITHIN DOTTED LINES SHEWN ON THE RIGHT. INEQUALITY FOR UNIMODAL, SYMMETRIC DISTRIBUTION.

a series of such inequalities in terms of the absolute moments of different orders, and for a slightly more general situation than the one given above. A proof of these inequalities is given in § V below. A proof for a still more general series of inequalities of the same type is given in Frechet (1937). It is also proved in § V that the Gauss-Winkler is the best possible inequality in the sense that, if only the absolute moment of one order is known, no improvement can be made on the inequality in terms of that moment.

It is possible to obtain a precise analogue of the Gauss-Winkler inequality in the simplest (and most important) case given above. It is convenient, because of symmetry, to change the notation slightly when discussing this inequality. If the mean of the unimodal symmetrical distribution is μ, and the mean range for samples of n, w_n, then the fraction of d falling in the interval $(\mu - tw_n, \mu + tw_n)$ is $2p_d(t)$ say. As with the inequality in Fig. 1, the mathematical expression is rather complicated, and it is simplest to use Fig. 2. If the ordinate of the point on the appropriate curve in Fig. 2 with abscissa t is $2p(t)$, then for any d

$$2p(t) < 2p_d(t).$$

The derivation of the equations of the limiting curves shown in Fig. 2 is given in § IV below. From the equations it can be shown that, for small n, there is not much difference numerically between the general inequality of § I and the more restricted inequality of this section, except for fairly small t. Table 2 gives a clear comparison. For selected $y = 1-p$, given in the left-hand column of the table, we obtain the corresponding t for the inequality of § I. For intervals of these lengths we obtain the corresponding $2p$; for the inequality of § II, assuming the same w_n. In the table are given the values of $2q = 1-2p$. Thus the figures in the body of the table can be compared directly with the values of y in the left-hand column.

Table 2. $2q = \dfrac{2}{R_n(y)}\left\{\dfrac{1+y^{n+1}-(1-y)^{n+1}}{n+1} - y^{n+1} - y(1-y)^n\right\}$

y \ n	2	3	4	5	6	7	8	9	10
0·45	0·327	0·327	0·309	0·285	0·259	0·236	0·215	0·196	0·179
0·40	0·311	0·311	0·295	0·273	0·250	0·229	0·210	0·193	0·177
0·35	0·287	0·287	0·273	0·255	0·236	0·218	0·202	0·187	0·173
0·30	0·257	0·257	0·246	0·232	0·217	0·203	0·190	0·177	0·166
0·25	0·222	0·222	0·214	0·203	0·193	0·183	0·173	0·164	0·155
0·225	0·203	0·203	0·196	0·188	0·179	0·171	0·162	0·155	0·147
0·20	0·183	0·183	0·178	0·171	0·164	0·157	0·151	0·144	0·138
0·175	0·163	0·163	0·158	0·153	0·148	0·142	0·137	0·132	0·127
0·15	0·141	0·141	0·138	0·134	0·130	0·126	0·122	0·119	0·115
0·125	0·119	0·119	0·117	0·114	0·111	0·109	0·106	0·103	0·101
0·10	0·096	0·096	0·095	0·093	0·091	0·090	0·088	0·086	0·085
0·075	0·073	0·073	0·072	0·071	0·070	0·069	0·068	0·067	0·066
0·05	0·049	0·049	0·049	0·048	0·048	0·047	0·047	0·047	0·046

As in § I, it may be desirable to have a more accurate estimate of $2p(t)$ than that obtained from the figure. Such an estimate can be obtained from Tables 1 and 2 as follows:

First, as in § I, find, for the given value of $2t = 1/R_n(y)$, the corresponding value of y, by harmonic interpolation in Table 1. Next, for the more restricted inequality with which we are now dealing, it is necessary to correct this value of y to a value $2q$, which can be found from Table 2 by direct interpolation.

Since $2p = 1-2q$, $2p$ can then be obtained immediately.

As an illustration, suppose that as in the example of § I, $n = 6$ and $2t = 1{\cdot}658$ (with the change of notation of this section). As before, the corresponding value of y is 0·143. Table 2 gives the value of $2q$ for $y = 0{\cdot}143$ as 0·125, giving 0·875 as the value of $2p$ required.

For large values of t, a fairly good approximation for the equations of the limiting curves is $2p(t) = 1-1/(2nt)$. This, remembering the change of notation, is the same approximation as that for the curves of Fig. 1. The advantage in assuming a unimodal symmetrical distribution lies in the fact that we deal with an interval in a known position relative to the distribution rather than in an unknown one.

Uses of these inequalities

When, in a quality control system, some property or dimension of a product from a process is being checked, the inspector will take small samples and often note only the mean and range of each Thus in time a very reliable estimate of the mean range of the process is available. The question then sometimes arises: If the tolerance interval for this process is

fixed at, say, l, how many rejects will be obtained? To answer this question the inspector would have to know the process distribution exactly, but with the help of these inequalities he could give a partial answer. First, if he knew that the distribution was symmetrical and unimodal, he could state that if the process mean, i.e. the mean of the distribution, were set accurately on the drawing mean, then the percentage of rejects would never be greater than a value obtained from the unimodal symmetrical inequality. Secondly, if the distribution was not unimodal and symmetrical, the inspector could state that, after practical experience had shown what was the best place to set the process mean and if this setting were held accurately, then the percentage of rejects produced would not be greater than a value obtained from the general inequality given here.

III. Derivation of the general inequality in terms of mean range

The distribution function

Any frequency distribution can be represented by its distribution function $F(x)$, which is defined as the fraction of the distribution falling on, and to the left of, the point x.

$F(x)$ will satisfy the following conditions:

(a) $F(x)$ is a monotonic increasing function,

(b) $\lim_{x \to -\infty} F(x) = 0$,

(c) $\lim_{x \to +\infty} F(x) = 1$,

and the inequality given below will apply to any distribution which can be represented by a function $F(x)$ satisfying these conditions, together with (d) below.

We will use the notation

$$\lim_{h \to 0} F(x+h) = F(x+0) \quad h \text{ positive,}$$
$$= F(x-0) \quad h \text{ negative.}$$

Then

(d) $F(x) = F(x+0)$.

From the conditions on F, it follows that $F(x) = F(x-0)$ except in an enumerable set of points.

The mean range

It can be shown that, for any distribution, if the mean range exists, it is given by

$$w_n = \int_{-\infty}^{\infty} R_n(F)\, dx \quad \text{(see Kendall, 1943),}$$

where $R_n(F) = 1 - F^n - (1-F)^n$.

Notice that, as F increases from 0 to $\frac{1}{2}$, $R_n(F)$ increases from 0 to $1 - \frac{1}{2}^{n-1}$, and as F increases from $\frac{1}{2}$ to 1, $R_n(F)$ decreases from $1 - \frac{1}{2}^{n-1}$ to 0.

The limiting curves

We will find the form of the limiting curves by finding the value of t for any value of p, $0 < p < 1$.

Suppose L is a fixed positive number and $L = tw_n$. Consider distributions satisfying the following conditions:

(i) There is an interval $(x', x'+L)$ such that

$$F(x'+L)-F(x'-0) = p,$$

i.e. a fraction p of the distribution lies in a closed interval of length L.

(ii) $$F(x+L)-F(x-0) \leqslant p \text{ for all } x.$$

For fixed L and p, we will find the lower bound of w_n for distributions satisfying conditions (i) and (ii). This will give an upper bound to t, and this will be the abscissa of the point on the limiting curve with ordinate p. For the upper bound of t is a monotonic increasing function of p for $0 \leqslant p < 1$. Consequently, for any t, no distribution can give a point below the curve.

In the first place, consider the case $p > \frac{1}{2}$.

Since $F(x-0) = F(x)$ except in a set of points which is enumerable, and therefore of measure zero,

$$\int_a^b R_n\{F(x)\}\,dx = \int_a^b R_n\{F(x-0)\}\,dx \text{ for any interval } (a, b).$$

If x' is a point satisfying condition (i) above, we can write

$$w_n = \int_{-\infty}^{x'} R_n\{F(x-0)\}\,dx + \int_{x'}^{\infty} R_n\{F(x)\}\,dx,$$

remembering $F(x'-0) \leqslant 1-p < \frac{1}{2}$.

The function F_1

Now we introduce a new distribution with distribution function F_1. Roughly speaking, this is obtained from F by compressing the distribution represented by F about its median. The formula for w_n shows that this reduces the mean range of a distribution. F_1 represents a finite distribution, and it has the property that, if $(x, x+L)$ lies entirely in the range of the distribution

$$F_1(x+L) = F_1(x)+p \text{ almost everywhere.}$$

This property enables us to find a lower bound to the mean range of F_1, and therefore to the mean range of F. An example shows that the bound obtained is the greatest lower bound.

We define the new function $F_1(x)$ as follows:

$$\begin{array}{lll} \text{If} & F(x+L)-p \leqslant 0 & F_1(x) = 0, \\ & F(x+L)-p \geqslant 0 \text{ and } x \leqslant x' & F_1(x) = F(x+L)-p, \\ & x' < x \leqslant x'+1 & F_1(x) = F(x), \\ & F(x-L-0)+p \leqslant 1 \text{ and } x \geqslant x'+L & F_1(x) = F(x-L-0)+p, \\ & F(x-L-0)+p \geqslant 1 & F_1(x) = 1. \end{array}$$

F_1 is uniquely defined at every point and is a monotonic increasing function.

By condition (ii) $$\int_{-\infty}^{x'} R_n\{F_1(x)\}\,dx \leqslant \int_{-\infty}^{x'} R_n\{F(x-0)\}\,dx,$$

since $F_1(x) \leqslant F(x-0)$ and $R_n(F)$ is an increasing function of F in this range.

Also by condition (ii) $$\int_{x'+L}^{\infty} R_n\{F_1(x)\}\,dx \leqslant \int_{x'+L}^{\infty} R_n\{F(x)\}\,dx,$$

since $F_1(x) \geqslant F(x)$ and $R_n(F)$ is a decreasing function of F in this range.

If w_n' is defined as $\int_{-\infty}^{\infty} R_n\{F_1(x)\}\,dx$, then

$$w_n' \leqslant w_n.$$

By a Dedekindian argument, there is a point x_0 such that $F_1 = 0$ for $x < x_0$, $F_1 > 0$ for $x \geqslant x_0$. Take x_0 as the new origin.

In the same way there is a point x_1 such that $F < 1$ for $x < x_1$, $F_1 = 1$ for $x \geqslant x_1$.

Define r by the equation
$$x_1 = L + r.$$

Now by the definition of $F_1(x)$, $F_1(x+L) - F_1(x) = p$ for $0 < x < r$, so for $x < r$, $F_1(x) < 1-p$, $x \geqslant r$, $F_1(x) \geqslant 1-p$.

For any $h > 0$, $F_1(L+h) \geqslant p$ and $F_1(r-h) < 1-p$, so $F_1(L+h) \geqslant p > 1-p \geqslant F_1(r-h)$.

Hence $L+h \geqslant r-h$ for all $h > 0$, so that $r \leqslant L$.

By the properties of F_1

$$w'_n = \int_0^{r+L} R_n(F_1)\,dx = \int_0^r \{R_n(F_1) + R_n(F_1+p)\}\,dx + \int_r^L R_n(F_1)\,dx.$$

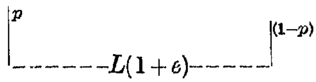
For
$$r \leqslant x \leqslant 1,\; 1-p \leqslant F_1 \leqslant p,\text{ so } R_n(F_1) \geqslant R_n(p),$$
$$0 \leqslant x \leqslant r,\; F_1 \leqslant 1-p,\text{ so } R_n(F_1) + R_n(F_1+p) \geqslant R_n(p),$$
so $w'_n \geqslant (L-r)\,R_n(p) + rR_n(p)$, i.e. $w'_n \geqslant LR_n(p)$, and therefore $w_n \geqslant LR_n(p)$.

Hence if $L = tw_n$, for given $p > \frac{1}{2}$ the upper bound of t is $L/R_n(p)$. The equations of the limiting curves are therefore, for $p > \frac{1}{2}$, $t > \dfrac{2^{n-1}}{2^{n-1}-1}$,

$$1 - (1-p)^n - p^n = t^{-1}.$$

Proof that the inequality found is the 'best possible'

The inequality is the best possible of its type. Consider the distribution



,

i.e.
$$\left.\begin{array}{ll} x < 0 & F = 0, \\ 0 \leqslant x < L(1+e) & F = p, \quad (e > 0) \\ L(1+e) \leqslant x & F = 1, \end{array}\right\}$$

For this distribution $w_n = L(1+e)\,R_n(p)$, so if $t = L/w_n$, by making e sufficiently small we can obtain a point as near the limiting curve as we please.

Derivation of equation of limiting curves for $p \leqslant \frac{1}{2}$

The equations of the limiting curves for the case $p \leqslant \frac{1}{2}$ are not of practical importance. Their derivation is similar to that for the case $p > \frac{1}{2}$, only it is slightly more complicated.

Suppose
$$\frac{1}{m+1} < p \leqslant \frac{1}{m} \quad (m = 2, 3, 4, \ldots).$$

A point x exists with the property that for $x < x_2$, $F < \frac{1}{2}$, and for $x \geqslant x_2$, $F \geqslant \frac{1}{2}$. Take x_2 as origin.

Define $F_1(x)$ as follows ($s = 1, 2, 3, \ldots$):
$$F_1(x) = F(x+sL) - sp, \quad -sL \leqslant x \leqslant -(s-1)\,L,$$
$$F_1(x) = F(x), \quad x \geqslant 0,$$
unless $F(x+sL) - sp < 0$ when $F_1(x) = 0$

$$w_n = \int_{-\infty}^{\infty} R_n(F)\,dx \geqslant \int_{-\infty}^{\infty} R_n(F_1)\,dx = w'_n.$$

Now define a function $F_2(x)$ as follows:

$$F_2(x) = F_1(x-sL)+sp, \quad (s-1)L \leqslant x \leqslant sL,$$
$$F_2(x) = F_1(x), \qquad x \leqslant 0,$$

unless $F_1(x-sL)+sp > 1$ when $\qquad F_2(x) = 1.$

Then $$w_n'' = \int_{-\infty}^{\infty} R_n(F_2)\,dx \leqslant w_n' \leqslant w_n.$$

A point x_0 exists such that $\qquad F_2(x) = 0,\ x < x_0,$

$$F_2(x) > 0,\ x \geqslant x_0.$$

A point x_1 exists such that $\qquad F_2(x) = 1,\ x \geqslant x_1,$

$$F_2(x) < 1,\ x < x_1.$$

Define r by $x_1 - x_0 = mL + r$. Take origin at x_0. Then $mL \leqslant x_1 - x_0 \leqslant (m+1)L$ as can be shown by considering $F_2(mL+h)$ and $F_2(mL-L+r-h)$ for $h>0$, as in the case $p>\frac{1}{2}$.

By the properties of $F_2(x)$

$$\begin{aligned} w_n'' &= \int_0^{mL+r} R_n(F_2)\,dx \\ &= \int_0^r \{R_n(F_2)+R_n(F_2+p)+\ldots+R_n(F_2+mp)\}\,dx \\ &\quad + \int_r^L \{R_n(F_2)+R_n(F_2+p)+\ldots+R_n(F_2+\overline{m-1}\,p)\}\,dx. \end{aligned}$$

For $\qquad 0 \leqslant x \leqslant r, \quad 0 < F_2 \leqslant 1-mp,$

so $$R_n(F_2)+R_n(F_2+p)+\ldots+R_n(F_2+mp) \geqslant R_n(p)+R_n(2p)+\ldots+R_n(mp).$$

For $\qquad r \leqslant x < L, \quad 1-mp \leqslant F_2 < p,$

so $$R_n(F_2)+R_n(F_2+p)+\ldots+R_n(F_2+\overline{m-1}\,p) \geqslant R_n(p)+R_n(2p)+\ldots+R_n(mp),$$

so $$\begin{aligned} w_n'' &\geqslant r\sum_{i=1}^m R_n(ip) + (L-r)\sum_{i=1}^m R_n(ip) \\ &= L\sum_{i=1}^m R_n(ip). \end{aligned}$$

The equation of the limiting curve is therefore

$$\sum_{i=1}^m R_n(ip) = t^{-1}$$

for $$\frac{1}{m+1} < p \leqslant \frac{1}{m},$$

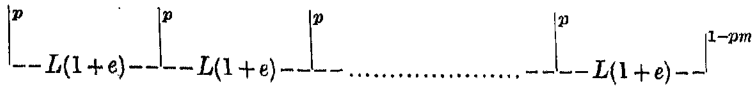
i.e. for $$\frac{1}{\sum_{i=1}^m R_n\left(\frac{i}{m+1}\right)} < t \leqslant \frac{1}{\sum_{i=1}^m R_n\left(\frac{i}{m}\right)}.$$

Note that $t = \dfrac{1}{\sum_{i=1}^m R_n\left(\frac{i}{m}\right)}$ is on the curve for all integral m.

As $m \to \infty$,
$$\frac{1}{m}\sum_{i=1}^{m} R_n\left(\frac{i}{m}\right) \to \int_0^1 R_n(x)\,dx > 0,$$
so
$$\frac{1}{\sum_{i=1}^{m} R_n\left(\frac{i}{m}\right)} \to 0,$$
and the curve starts at $t = 0$, $p = 0$.

Again the inequality found is the 'best possible'

The inequality is the best possible of its kind. For consider the frequency distribution



where
$$\begin{aligned}
F &= 0, && x < 0,\\
F &= p, && 0 \leqslant x < L(1+e),\\
F &= 2p, && L(1+e) \leqslant x < 2L(1+e),\\
&\cdots\cdots && \\
F &= mp, && (m-1)\,L(1+e) \leqslant x < mL(1+e),\\
F &= 1, && mL(1+e) \leqslant x.
\end{aligned}$$

By the formula for mean range given above
$$w_n = \int_{-\infty}^{\infty} R_n(F)\,dx = L(1+e)\sum_{i=1}^{m} R_n(ip).$$
so, for this distribution, if $t = L/w_n$
$$t^{-1} = (1+e)\sum_{i=1}^{m} R_n(ip),$$
and this point can, for sufficiently small e, be as near the limiting curve as we please.

IV. Derivation of the inequality for unimodal symmetrical distributions

There will be some differences between the notation used in this section and that of the preceding section.

The inequality gives the lower bound to the fraction, $2p$, of a distribution, falling in an interval $(\mu - tw_n, \mu + tw_n)$, where μ is the mean of the distribution. In fact what we do is to find the parametric equations to the curves of Fig. 2. As in § III, we find the upper bound for t for a given p $(0 < p < 1)$. We thus obtain a curve in the (t, p) plane for each n. Then the same curve gives the lower bound to p for each t, since it is monotonic.

Suppose $L = tw_n$ and L is fixed. Consider unimodal symmetrical distributions such that:

(*a*) their means are at the centre of an interval of length $2L$,

(*b*) exactly $2p$ of the distribution lies in this interval.

We must find the minimum of w_n for distributions satisfying these conditions.

Since the distributions considered are unimodal and symmetrical, they can only have a saltus (where there is a finite fraction of the distribution concentrated at a single point) at the mean. Take the origin at the mean. Elsewhere $f(x) = F'(x)$ is finite. We only need consider positive x.

Now suppose h is a number less than or equal to p/L. Consider, for the moment, only distributions satisfying $f(L) = h$. In order that $\int_L^\infty R_n(F)\,dx$ should be a minimum, $F(x)$ must be as large as possible for all x exceeding L, but $F(L) = \frac{1}{2}+p$, which is fixed. Clearly, then, the distribution maximizing F, and so minimizing $R_n(F)$, is rectangular for x greater than L, i.e. if $p+q = \frac{1}{2}$,

$$f(x) = h \quad \text{for} \quad L \leqslant x \leqslant \frac{q}{h}+L,$$

$$f(x) = 0 \quad \text{for} \quad \frac{q}{h}+L < x.$$

Consider now $0 < x \leqslant L$. As x decreases from L to 0, F must decrease as little as possible. Consequently to minimize $R_n(F)$ the distribution must be rectangular in this interval, and finally there must be a saltus at the origin of amount $2(p-hL)$.

It remains to find the value of h which minimizes w_n for the type of distribution given.

Move the origin to the start of this distribution:

$$\begin{aligned} w_n &= 2\int_0^{(q+hL)/h} \{1-F^n-(1-F)^n\}\,dx \\ &= 2\int_0^{(q/h)+L} \{1-h^n x^n-(1-hx)^n\}\,dx \\ &= 2\left\{\frac{q}{h}+L-\frac{(q+hL)^{n+1}}{h(n+1)}+\frac{1}{h(n+1)}(1-q-Lh)-\frac{1}{h(n+1)}\right\}. \end{aligned}$$

Let $hL = x$, $\dfrac{L}{w_n} = t$,

$$\frac{1}{2t} = \left[1+\frac{q}{x}-\frac{1}{(n+1)x}\{1+(q+x)^{n+1}-(1-q-x)^{n+1}\}\right], \tag{1}$$

so $$\frac{x^2}{2}\frac{d}{dx}(t^{-1}) = -q+\frac{1}{n+1}\{1+(q+x)^{n+1}-(1-q-x)^{n+1}\}-x\{(q+x)^n+(1-q-x)^n\}. \tag{2}$$

For minimum $w_n = \dfrac{1}{t}$ we must have $\dfrac{d}{dx}\left(\dfrac{1}{t}\right) = 0$.

Now $$\frac{d}{dx}\left\{\frac{x^2}{2}\frac{d}{dx}(t^{-1})\right\} = nx\{(1-q-x)^{n-1}-(q+x)^{n-1}\} > 0,$$

so $\dfrac{x^2}{2}\dfrac{d}{dx}(t^{-1})$ is a monotonic increasing function.

Also when $x = 0$ $$\frac{x^2}{2}\frac{d}{dx}(t^{-1}) = \frac{1}{n+1}-q-\frac{1}{n+1}\{(1-q)^{n+1}-q^{n+1}\}. \tag{3}$$

When $q = 0$ the right-hand side of (3) is 0, and its differential coefficient with respect to q is $-R_n(q) \leqslant 0$.

Hence $$\frac{1}{n+1}-q-\frac{1}{n+1}\{(1-q)^{n+1}-q^{n+1}\} \leqslant 0,$$

so $\dfrac{x^2}{2}\dfrac{d}{dx}(t^{-1})$ is negative when x is small, since $q \neq 0$.

Since $\dfrac{x^2}{2}\dfrac{d}{dx}(t^{-1})$ and $\dfrac{d}{dx}(t^{-1})$ have the same zeros if $x \neq 0$, $\dfrac{d}{dx}(t^{-1})$ has either one or no zeros in $(0, p)$.

If $\frac{x^2}{2}\frac{d}{dx}(t^{-1})$ is negative when $x = p$, the minimum of $\frac{1}{t}$ is at $x = p$. Hence if

$$-q+\frac{1}{n+1}-\frac{p}{2^{n-1}}<0,$$

i.e. if $p<\frac{n-1}{n+1}\frac{2^{n-2}}{2^{n-1}-1}$, the equation of the limiting curve will be $\frac{1}{2t}=1+\frac{q}{p}-\frac{1}{(n+1)p}$, i.e.

$$p=\frac{n-1}{n+1}t. \tag{4}$$

If $p\geqslant\frac{n-1}{n+1}\frac{2^{n-2}}{2^{n-1}-1}$ then $\frac{d}{dx}(t^{-1})$ will have a zero in $(0, p)$. To obtain the equation of the limiting curve in this case we have to combine the two equations

$$\frac{d}{dx}(t^{-1})=0,$$

and
$$\frac{1}{2t}=\left[1+\frac{q}{x}-\frac{1}{(n+1)x}\{1+(q+x)^{n+1}-(1-q-x)^{n+1}\}\right]. \tag{1 bis}$$

Put $q+x=y$. Then $\frac{d}{dx}(t^{-1})=0$ gives

$$-q+\frac{1}{n+1}\{1+y^{n+1}-(1-y)^{n+1}\}=(y-q)\{R_n(y)-1\}. \tag{5}$$

From (1) and (2) the parametric equations of the limiting curves for the case

$$t=\frac{2^{n-2}}{2^{n-1}-1},\quad p=\frac{n-1}{n+1}\frac{2^{n-2}}{2^{n-1}-1}$$

are
$$\frac{1}{2t}=R_n(y)$$

and
$$q=\frac{1}{R_n(y)}\left\{\frac{1+y^{n+1}-(1-y)^{n+1}}{n+1}-y^{n+1}-y(1-y)^n\right\}.$$

If y is given any value between 0 and $\frac{1}{2}$, we obtain a point on the curve.

As $y\to 0$, $2q-y$ is $0(y^2)$ so that for small y the limiting curve gives values not far from the curve:

$$\frac{1}{2t}=R_n(2q).$$

Since we have actually found distributions giving points on the limiting curve, no better inequality of this type can be found.

Computation

Table 2 was calculated from the formula

$$q=\frac{1}{(n+1)(1-2y)R_n(y)}\{R_{n+1}(y)+(n-1-2ny)yR_n(y)-y(n-1-2ny)-2y\}.$$

$R_n(y)$ can be calculated from the recurrence relation

$$R_{n+1}(y)=R_n(y)+y(1-y)\{1-R_{n-1}(y)\}.$$

V. A derivation of the Gauss-Winkler inequalities

A method similar to that used in § III above can be applied to prove the Gauss-Winkler inequalities. Suppose $f(x) = F'(x)$ and $f_1(x) = f(x)+f(-x)$. The inequality holds for all distributions such that $f_1(x)$ is a decreasing function of x.

If λ_r^r is the rth absolute moment about the origin

$$\lambda_r^r = \int_0^\infty f_1(x)\, x^r\, dx.$$

The inequalities give the lower bound to the fraction $2p$ of a distribution lying in the interval $(-t\lambda_r, t\lambda_r)$.

Consider distributions such that exactly $2p$ of the distribution lies in the interval $(-L, L)$. We will find the lower bound of λ_r for such distributions. L and p are considered as fixed.

Consider first distributions satisfying the additional condition

$$f_1(L) = h \quad \left(h < \frac{2p}{L}\right).$$

If $2q = 1-2p$, $\int_L^\infty f_1(x)\, dx = 2q$ a constant, so to minimize $\int_L^\infty x^r f_1(x)\, dx, f_1(x)$ must be constant for $x \geqslant L$.

Thus $$f_1 = h \quad \text{for} \quad L \leqslant x \leqslant \frac{2q}{h}+L, \quad f_1 = 0 \quad \text{for} \quad L+\frac{2q}{h} < x.$$

To minimize $\int_0^L x^r f_1(x)\, dx$, $f_1(x)$ should be constant, therefore $f_1 = h$ for $0 < x \leqslant L$, and as $h < \frac{2p}{L}$ there is a saltus in F of amount $2p - Lh$ at the origin.

Now we must find h to minimize $\int_0^\infty f_1(x)\, x^r\, dx$:

$$\lambda_r^r = \int_0^{L+(2q/h)} hx^r\, dx = \frac{h}{r+1}\left(L+\frac{2q}{h}\right)^{r+1},$$

so $$\frac{d\lambda_r^r}{dh} = \frac{\left(L+\frac{2q}{h}\right)^r}{r+1}\left\{L+\frac{2q}{h} - 2(r+1)\frac{q}{h}\right\}$$

and $\frac{d^2\lambda_r^r}{dh^2}$ is positive, so λ_r^r is a minimum when $h = \frac{2rq}{L}$.

Since in any case $h \leqslant \frac{2p}{L}$ the inequality takes two different forms. If $\frac{2rq}{L} \geqslant \frac{2p}{L}$, i.e. if $2p \leqslant \frac{r}{r+1}$, the minimum is obtained when $h = \frac{2p}{L}$. In this case minimum $\lambda_r^r = \frac{L^r}{(r+1)(2p)^r}$, since $2p+2q = 1$.

If $2p \geqslant \frac{r}{r+1}$, minimum $\lambda_r^r = 2q\frac{r^r}{(r+1)^r L^r}$, so if $\frac{L}{\lambda r} = t$, the equations of the limiting curves will be:

$$\text{for} \quad t = \frac{r}{(r+1)^{1-1/r}}, \quad 2q = 1 - \frac{t}{(r+1)^{1/r}},$$

$$t = \frac{r}{(r+1)^{1-1/r}}, \quad 2q = \frac{r^r}{(r+1)^r t^r}.$$

In particular, if the distribution is symmetrical, the mode coincides with the mean. If the origin is at the mean, $\lambda^2 = \sigma^2$, and we obtain the familiar inequalities

$$2q \leqslant 1 - \frac{t}{\sqrt{3}}, \text{ if } t \leqslant \frac{2}{\sqrt{3}},$$

$$2q \leqslant \frac{4}{9t^2}, \quad \text{if } t \geqslant \frac{2}{\sqrt{3}}.$$

Since during the course of the proof, we found distributions satisfying the equality, no better inequality is possible.

SUMMARY

Two inequalities are found in terms of mean range of samples of n ($n = 2, 3, 4, \ldots$). The first is true, as is Tchebycheff's, for any frequency distribution whatsoever. The second holds for any unimodal symmetrical distribution. Both are shown to be the best possible of their type. Diagrams are given to facilitate the use of the inequalities for the cases $n = 2, 3, \ldots, 9, 10$. A derivation is also given of the Gauss-Winkler inequalities analogous to that used for inequalities in terms of mean range.

This paper was written in the course of my work at the Ministry of Supply (S.R. 17). My attention was drawn to this problem by a note Q.C./R/11 (issued by S.R. 17) by G. A. Barnard, to whom I offer my thanks.

REFERENCES

FRECHET, M. (1937). *Recherches Théoriques Modernes sur la Théorie des Probabilités.* Paris.

KENDALL, M. G. (1943). *Advanced Theory of Statistics.* London: Charles Griffin.

TABLES FOR TESTING THE HOMOGENEITY OF A SET OF ESTIMATED VARIANCES

COMPUTED BY CATHERINE M. THOMPSON AND MAXINE MERRINGTON

PREFATORY NOTE BY H. O. HARTLEY AND E. S. PEARSON

1. HISTORICAL NOTE

The statistical analysis of data often leads to the calculation of a number of estimated variances of which it is desirable to test the homogeneity. The present tables have been computed to facilitate the application of this test.

As far as we are aware, the original test of this nature was that obtained by J. Neyman & E. S. Pearson (1931), who suggested the use of a criterion L_1 which was the ratio of a weighted geometric to a weighted arithmetic mean of the mean squares from which the variances were estimated. On the assumption that variation followed the normal law, these authors: (a) gave the sampling moments of L_1 if the hypothesis of equal sampling variances was true; (b) showed that in the case of large samples $-N\log L_1$ was distributed as χ^2 with $k-1$ degrees of freedom (where N was the total number of observations and k the number of separate estimates of variance); (c) suggested a method of calculating approximate probability levels for L_1 in the case of small samples.

Following this line of attack, other contributions were made by B. L. Welch (1935, 1936), who showed how L_1 could be generalized and the weighting, chosen for the different sums of squares, modified; by P. P. N. Nayer (1936), who computed tables of probability levels of L_1 for the case of equal samples; and by U. S. Nair (1938), who investigated the form of the true distribution of L_1.

Meanwhile M. S. Bartlett (1937), approaching from another angle, suggested an analogous test in which the sums of squares were weighted with their appropriate degrees of freedom instead of with the number of observations as in the Neyman-Pearson criterion. Thus if s_t^2 is the usual unbiased estimate of σ_t^2, based on a sum of squares having ν_t degrees of freedom, and there are k of these estimates from independent sets of observations, Bartlett took as his test function

$$-2\log\mu = \mathrm{N}\log\left\{\sum_{t=1}^{k}(\nu_t s_t^2)/\mathrm{N}\right\} - \sum_{t=1}^{k}(\nu_t \log s_t^2), \tag{1}$$

$$\mathrm{N} = \sum_{t=1}^{k}(\nu_t), \tag{2}$$

and natural logarithms to base e are used. Provided that none of the degrees of freedom ν_t are too small, $-2\log\mu$ is distributed approximately as χ^2 with $k-1$ degrees of freedom if the null hypothesis is true, i.e. if the σ_t^2 $(t = 1, 2, \ldots, k)$ have a common value. For small samples, Bartlett introduced the corrective factor

$$C = 1 + \frac{1}{3(k-1)}\left\{\sum_t \frac{1}{\nu_t} - \frac{1}{\mathrm{N}}\right\}, \tag{3}$$

and showed that the quantity $-(2\log\mu)/C$ followed approximately the same χ^2 distribution law.

A comparison of these tests was given by D. J. Bishop & U. S. Nair (1939), who showed that even using the C correction the χ^2 approximation is not altogether satisfactory if some of the degrees of freedom, ν_t, are 1, 2 or 3.

In a later paper H. O. Hartley (1940) derived another method of approximating to the distribution of Bartlett's $-2\log\mu$, in which the probability integral is represented as a weighted mean of χ^2 integrals. This approximation is sufficiently accurate to allow the degrees of freedom to drop to 2; even if some estimates of variance based on 1 degree of freedom are included among the k values, the approximation is still fair.

The tables published below are based on Hartley's approximation; in presenting them to statisticians for general use it is hoped to render this test both more convenient and more accurate.

2. General scheme of the new tables

It is supposed that the data fall into k groups within each of which a random variable x is normally distributed with variance σ_t^2 $(t = 1, 2, \ldots, k)$. s_t^2 is the usual, unbiased sample estimate of σ_t^2 based on a sum of squares having ν_t degrees of freedom. The question at issue is whether the data are heterogeneous as to variance or whether they are consistent with the hypothesis that all σ_t^2 $(t = 1, 2, \ldots, k)$ have a common, if unknown, value.

The test is carried out by calculating

$$M = \mathrm{N}\log_e\left\{\sum_t(\nu_t s_t^2)/\mathrm{N}\right\} - \sum_t(\nu_t \log_e s_t^2),^* \tag{4}$$

where

$$\mathrm{N} = \sum_t(\nu_t).$$

Hartley (1940) has shown that if there is a common variance, the probability distribution of M can be closely described in terms of three parameters, namely k, c_1 and c_3, where

$$c_1 = \sum_t\left(\frac{1}{\nu_t}\right) - \frac{1}{\mathrm{N}}, \tag{5}$$

$$c_3 = \sum_t \frac{1}{\nu_t^3} - \frac{1}{\mathrm{N}^3}. \tag{6}$$

Tables 1 and 2 below enable the 5 % and 1 % significance levels of M to be obtained. They are tables of double entry for k and c_1; for each combination of these two quantities it will be seen that there are two entries denoted by (a) and (b). These are approximately maximum and minimum values of the true percentage point which will normally have an intermediate value, dependent on c_3. Provided the degrees of freedom are not very unequal, the correct value of M will be close to the entry opposite (a).

The tables have been arranged to make their use as simple as possible. If in the table of 5 % points, say, all entries in the lines for the appropriate k are greater than the value of M derived from the data, this value is not significant at the 5 % level. On the other hand, if the calculated M is larger than all entries for that k, then M is significant at the 5 % level. In neither case is it necessary to calculate c_1 or c_3. When, however, M falls within the range of values shown in the lines for the particular k, it is necessary to calculate c_1 from equation (5). Knowing this value, it will usually be possible to form an opinion on the significance

* We propose to use the single letter M in place of Bartlett's $-2\log_e\mu$.

of M without proceeding to the calculation of c_3 needed for interpolation between the entries (*a*) and (*b*). A description of this more refined procedure is, however, given in § 4 below.

It will be noted that the entries in the tables under $c_1 = 0$ are simply the 5 % and 1 % probability levels of χ^2 with $k-1$ degrees of freedom, this being the limiting form approached when all the ν_t are large. On the other hand, for a given k, c_1 has its maximum value when all ν_t are unity and therefore $c_1 = k - 1/k$.*

3. An illustrative example

The use of the table is best demonstrated in terms of an example. Below is shown (col. 3) a set of ten estimates of variance, calculated from ten samples of weight records of schoolboys of similar age, but from different forms. It is desired to test whether there are any real 'form differences' in the weight dispersion of the boys. To this end we set out the calculations of M as shown below:

(1) Form no. t	(2) No. of boys n_t	(3) Weight variance s_t^2 (lb.^{2})	(4) ν_t	(5) $\log_e s_t^2$	(6) $\nu_t \log_e s_t^2$	(7) $1/\nu_t$
1	10	51	9	3·93	35·4	0·111
2	15	78	14	4·36	61·0	0·071
3	21	91	20	4·51	90·2	0·050
4	23	52	22	3·95	86·9	0·045
5	15	101	14	4·62	64·7	0·071
6	11	36	10	3·58	35·8	0·100
7	31	41	30	3·71	111·3	0·033
8	15	76	14	4·33	60·6	0·071
9	3	64	2	4·16	8·3	0·500
10	6	93	5	4·53	22·6	0·200
Totals	150		140 ($=N$)		576·8	1·252

We obtain further:

$$\Sigma\nu_t s_t^2 = 9176, \quad \Sigma(\nu_t s_t^2)/\mathrm{N} = 65{\cdot}54, \quad \log_e\{\Sigma\nu_t s_t^2/\mathrm{N}\} = 4{\cdot}183.$$

Hence $$M = 140 \times 4{\cdot}183 - 576{\cdot}8 = 8{\cdot}8.$$

The observed value of the 'variance dispersion', M, is therefore 8·8. This has to be compared with the appropriate tabulated 5 % (or 1 %) point. It is seen from Table 1 that all entries opposite $k = 10$ are greater than 8·8. Without further calculation it may therefore be concluded that M is not significant at the 5 % level, and we may infer that no real differences are indicated in the weight dispersion among the ten forms of schoolboys.

Had the observed value of M been 18·8 (instead of 8·8), the decision as to its significance would not have been obvious, since some of the 5 % points tabulated in the lines for $k = 10$ are smaller than 18·8.† It is now necessary to calculate c_1, defined in equation (5). Using the reciprocals of ν_t given in col. 7 of the table above, it is found that

$$c_1 = 1{\cdot}25.$$

* Actually, the last entry in each line has been computed from the approximating function, putting $c_1 = k$.

† Reference to Table 2 shows, however, that M cannot be significant at the 1 % level.

Since the percentage points (a) and (b) for $k = 10$ and both $c_1 = 1\cdot0$ and $1\cdot5$ are less than 18·8, we can say that M would now be significant at the 5 % level.

Had the data given a value of 17·6 for M which lies between the four appropriate tabled entries, it would normally have satisfied our purpose merely to note that M was on the border line of 5 % significance. If more precise information is needed, it will be necessary to proceed further by calculating c_3 and interpolating as indicated in the following section.

4. Definition of the table entries (a) and (b); the use of c_3

It can be shown that, for a given k and c_1, the range of c_3 is (to a first order of approximation)

$$c_3(a) = c_1^3/k^2 < c_3 < c_1 = c_3(b). \tag{7}$$

The lower bound is approached when all values of ν_t are equal and the upper bound when j, say, of the ν_t are each equal to unity and the $k-j$ remaining values all tend to infinity. In Tables 1 and 2 the entry for the percentage point opposite (a) is that for $c_3 = c_3(a)$; that opposite (b) is for $c_3 = c_3(b)$. It will be seen that, at any rate throughout the tabulated range of values, the entry (a) is greater than or equal to (b). In using the former, therefore, we shall in rare cases fail to detect the significance of M.

If interpolation for c_3 is decided on, use may be made of the auxiliary Table 3. This gives for all the marginal entries k and c_1 of Tables 1 and 2, the two quantities

$$C = c_3(a) = c_1^3/k^2, \quad \Delta C = c_3(b) - c_3(a) = c_1 - c_1^3/k^2. \tag{8}$$

The procedure would then be first to interpolate linearly in the two nearest c_1 columns between the two percentage points (a) and (b), using the formula:

Percentage point corresponding to c_3

$$\approx \frac{1}{\Delta C}\{(c_1 - c_3) \times \text{entry}\,(a) + (c_3 - C) \times \text{entry}\,(b)\}, \tag{9}$$

and then interpolating to the correct value of c_1.

Example. Suppose that $k = 10$ and the degrees of freedom are the ten values of ν_t given in col. 4 of the illustrative data tabled above. Here

$$c_1 = 1\cdot25, \quad c_3 = 0\cdot14.$$

For the interpolation process, we need the following entries:

		$c_1 = 1\cdot0$	$c_1 = 1\cdot5$
From Table 3	C	0·010	0·034
	ΔC	0·990	1·466
From Table 1	Entry (a)	17·54	17·83
	Entry (b)	17·17	17·29

Hence the 5 % point corresponding to $c_1 = 1\cdot0$, $c_3 = 0\cdot14$ is approximately, from equation (9)

$$\frac{1}{0\cdot99}\{0\cdot86 \times 17\cdot54 + 0\cdot13 \times 17\cdot17\} = 17\cdot49.$$

The 5 % point corresponding to $c_1 = 1\cdot5$ and $c_3 = 0\cdot14$ will be approximately

$$\frac{1}{1\cdot47}\{1\cdot36 \times 17\cdot83 + 0\cdot11 \times 17\cdot29\} = 17\cdot79.$$

Interpolating between these two values for $c_1 = 1\cdot25$, we find finally a 5 % point for M at 17·64. It will be seen that this value differs very little from that of 17·68 obtained by using the (a) entries only.

5. Accuracy of the approximation

To test the accuracy of Hartley's approximation, we may compare the present tables with the values worked out by Bishop & Nair (1939). Some of these values were calculated from

Nair's (1938) exact expansion applicable to the special case where all ν_t are equal; some were obtained by fitting a type I curve to the distribution of L_1 using a formula for its moments given by Welch (1936).

We deal first with the special case of $\nu_t = \nu$ for all t. In this case the parameter c_3 is very close to the value c_1^3/k^2 which is that one used for the percentage points (a). The comparisons are summarized in the table below:

k	ν	c_1	c_3	N	5 % points for M		1 % points for M	
					Bishop & Nair	Hartley	Bishop & Nair	Hartley
3	2	1·33	0·37	6	7·11*	7·05	10·74*	10·57
	3	0·89	0·11	9	6·80†	6·79	10·43†	10·32
	4	0·67	0·05	12	6·62*	6·61	10·13*	10·10
	9	0·30	0·00	27	6·30†	6·28	9·67†	9·64
5	2	2·40	0·62	10	11·09*	11·01	15·32*	15·15
	3	1·60	0·18	15	10·67†	10·62	14·91†	14·76
	4	1·20	0·08	20	10·38*	10·37	14·47*	14·46
	9	0·53	0·01	45	9·93†	9·90	13·86†	13·84
10	2	4·95	1·25	20	19·62*	19·45	24·90*	24·65
	3	3·30	0·37	30	18·82†	18·79	24·09†	23·97
	4	2·48	0·16	40	18·42*	18·38	23·34*	23·49
	9	1·10	0·01	90	17·64†	17·60	22·48†	22·53

* Calculated from Nair's exact distribution. † Calculated by fitting type I curve to L_1.

The second decimal of the results calculated from Bishop & Nair's three-figure table is not always reliable. In view of this, the agreement for $\nu \geqslant 3$ is very good and that for $\nu = 2$ is certainly better than that with Bartlett's approximation, given in Table 1*b* of Bishop & Nair's paper.

For $\nu = 1$ the approximation breaks down; for example, for $k = 4$, $\nu = 1$ we have:

	5 % point	1 % point
Hartley's approximation	9·0	11·8
Nair's expansion	10·0	14·1

Next, we may make a few comparisons for the case of five estimates of variance having unequal degrees of freedom. In this general case an exact answer is no longer available for comparison and Bishop & Nair's values are, throughout, those obtained by fitting a type I curve to the distribution of L_1'. The comparisons are summarized in the table below:

ν_1	ν_2	ν_3	ν_4	ν_5	N	c_1	c_3	5 % point		1 % point	
								Bishop & Nair	Hartley	Bishop & Nair	Hartley
6	6	4	2	2	20	1·53	0·27	10·59	10·54	14·80	14·62
16	16	9	2	2	45	1·21	0·25	10·35	10·30	14·46	14·31
5	5	4	3	3	20	1·27	0·11	10·43	10·41	14·59	14·51
14	14	9	4	4	45	0·73	0·03	10·05	10·04	14·05	14·03

Again we see that where all degrees of freedom ν_t are greater than or equal to 3 the approximation is very good; where some of the degrees of freedom are as small as 2, the approximation is still adequate.

It must be noted that, throughout, the approximation has a *systematic* bias in that the values are consistently smaller than the exact ones or those obtained by fitting a type I curve. It is because of this systematic bias that the percentage point tabulated under (*a*) is sometimes actually nearer to the exact value than the one obtained by interpolation between the percentage point (*a*) and (*b*).

The question of whether linear interpolation between the percentage points is justified is not important where the systematic bias in the approximation is large. It will be noted that for all $\nu_t \geqslant 4$, when the approximation is expected to yield good results, linear interpolation between (*a*) and (*b*) gives the correct answer to about two-decimal accuracy. However, in these cases the interpolate is near to the percentage point (*a*), so that any second order term in the interpolation formula would have a small effect in any case.

6. THE CALCULATION OF THE PRESENT TABLES

The calculation of the present tables has been carried out according to formula (20), given by Hartley (1940). The values of the probability integral of χ^2 ($P_\kappa(\chi^2)$) were obtained from Table 12 of *Tables for Statisticians and Biometricians*, vol. 1 (1930, 3rd ed.). It was found necessary, however, to extend these tables beyond their present range of both χ^2 and n. This was done with the help of Molina's tables (1942), using the identity relation between the Poisson distribution and the χ^2 integral. These extended tables are available in manuscript at the Department of Statistics, University College.

We should like to record our appreciation of the extensive work undertaken by Miss Catherine M. Thompson (now Mrs Grylls) and Mrs Maxine Merrington in computing the tables.

REFERENCES

BARTLETT, M. S. (1937). *Proc. Roy. Soc.* A, **160**, 268.
BISHOP, D. J. & NAIR, U. S. (1939). *J.R. Statist. Soc. Suppl.* **6**, 89.
HARTLEY, H. O. (1940). *Biometrika*, **31**, 249.
MOLINA, E. C. (1942). *Poisson's Exponential Binomial Limit.* D. van Nostrand, New York.
NAIR, U. S. (1938). *Biometrika*, **30**, 274.
NAYER, P. P. N. (1936). *Statist. Res. Mem.* **1**, 38.
NEYMAN, J. & PEARSON, E. S. (1931). *Bull. int. Acad. Cracovie*, A, p. 460.
WELCH, B. L. (1935). *Biometrika*, **27**, 145.
WELCH, B. L. (1936). *Statist. Res. Mem.* **1**, 1.

Table 1. *M distribution*: 5 % *points*

k \ c_1		0·0	0·5	1·0	1·5	2·0	2·5	3·0	3·5	4·0	4·5	5·0	6·0	7·0	8·0	9·0	10·0	12·0	14·0
3	(a)	5·99	6·47	6·89	7·20	7·38	7·39	7·22	—	—	—	—	—	—	—	—	—	—	—
	(b)	5·99	6·22	6·43	6·64	6·84	7·03	7·22	—	—	—	—	—	—	—	—	—	—	—
4	(a)	7·81	8·24	8·63	8·96	9·21	9·38	9·43	9·37	9·18	—	—	—	—	—	—	—	—	—
	(b)	7·81	8·00	8·17	8·35	8·52	8·69	8·85	9·02	9·18	—	—	—	—	—	—	—	—	—
5	(a)	9·49	9·88	10·24	10·57	10·86	11·08	11·24	11·32	11·31	11·21	11·02	—	—	—	—	—	—	—
	(b)	9·49	9·65	9·80	9·96	10·11	10·27	10·42	10·57	10·72	10·87	11·02	—	—	—	—	—	—	—
6	(a)	11·07	11·43	11·78	12·11	12·40	12·65	12·86	13·01	13·11	13·14	13·10	12·78	—	—	—	—	—	—
	(b)	11·07	11·22	11·36	11·51	11·65	11·79	11·94	12·08	12·22	12·36	12·50	12·78	—	—	—	—	—	—
7	(a)	12·59	12·94	13·27	13·59	13·88	14·15	14·38	14·58	14·73	14·83	14·88	14·81	14·49	—	—	—	—	—
	(b)	12·59	12·73	12·87	13·00	13·14	13·27	13·41	13·55	13·68	13·82	13·95	14·22	14·49	—	—	—	—	—
8	(a)	14·07	14·40	14·72	15·03	15·32	15·60	15·84	16·06	16·25	16·40	16·51	16·60	16·49	16·16	—	—	—	—
	(b)	14·07	14·20	14·33	14·46	14·59	14·72	14·85	14·98	15·11	15·25	15·38	15·64	15·90	16·16	—	—	—	—
9	(a)	15·51	15·83	16·14	16·44	16·73	17·01	17·26	17·49	17·70	17·88	18·03	18·22	18·26	18·12	17·79	—	—	—
	(b)	15·51	15·63	15·76	15·89	16·02	16·14	16·27	16·40	16·52	16·65	16·78	17·03	17·29	17·54	17·79	—	—	—
10	(a)	16·92	17·23	17·54	17·83	18·12	18·39	18·65	18·89	19·11	19·31	19·48	19·75	19·89	19·89	19·73	19·40	—	—
	(b)	16·92	17·04	17·17	17·29	17·41	17·54	17·66	17·79	17·91	18·04	18·16	18·41	18·66	18·91	19·16	19·40	—	—
11	(a)	18·31	18·61	18·91	19·20	19·48	19·76	20·02	20·26	20·49	20·70	20·89	21·21	21·42	21·52	21·49	21·32	—	—
	(b)	18·31	18·43	18·55	18·67	18·79	18·91	19·04	19·16	19·28	19·40	19·52	19·77	20·01	20·26	20·50	20·75	—	—
12	(a)	19·68	19·97	20·26	20·55	20·83	21·10	21·36	21·61	21·84	22·06	22·27	22·62	22·88	23·06	23·12	23·07	22·56	—
	(b)	19·68	19·79	19·91	20·03	20·15	20·27	20·39	20·51	20·63	20·75	20·87	21·12	21·36	21·60	21·84	22·08	22·56	—
13	(a)	21·03	21·32	21·60	21·89	22·16	22·43	22·69	22·94	23·18	23·40	23·62	23·99	24·30	24·53	24·66	24·70	24·44	—
	(b)	21·03	21·14	21·26	21·38	21·50	21·62	21·74	21·85	21·97	22·09	22·21	22·45	22·69	22·92	23·16	23·40	23·88	—
14	(a)	22·36	22·65	22·93	23·21	23·48	23·75	24·01	24·26	24·50	24·73	24·95	25·34	25·68	25·95	26·14	26·25	26·17	25·66
	(b)	22·36	22·48	22·60	22·71	22·83	22·95	23·06	23·18	23·30	23·42	23·53	23·77	24·00	24·24	24·48	24·71	25·19	25·66
15	(a)	23·68	23·97	24·24	24·52	24·79	25·05	25·31	25·56	25·80	26·04	26·26	26·67	27·03	27·33	27·56	27·73	27·80	27·50
	(b)	23·68	23·80	23·92	24·03	24·15	24·26	24·38	24·50	24·61	24·73	24·85	25·08	25·31	25·55	25·78	26·01	26·48	26·95

$M = N \log_e \left\{ \sum_{t=1}^{k} (\nu_t s_t^2)/N \right\} - \sum_{t=1}^{k} (\nu_t \log_e s_t^2)$ $N = \sum_{t=1}^{k} \nu_t$ $c_1 = \sum_{t=1}^{k} \frac{1}{\nu_t} - \frac{1}{N}$ N.B. $\log_e x = 2{\cdot}3026 \log_{10} x$.

Table 2. *M distribution*: 1 % *points*

k \ c_1		0·0	0·5	1·0	1·5	2·0	2·5	3·0	3·5	4·0	4·5	5·0	6·0	7·0	8·0	9·0	10·0	12·0	14·0
3	(a)	9·21	9·92	10·47	10·78	10·81	10·50	9·83	—	—	—	—	—	—	—	—	—	—	—
	(b)	9·21	9·29	9·38	9·48	9·59	9·71	9·83	—	—	—	—	—	—	—	—	—	—	—
4	(a)	11·34	11·95	12·46	12·86	13·11	13·18	13·03	12·65	12·03	—	—	—	—	—	—	—	—	—
	(b)	11·34	11·40	11·46	11·54	11·63	11·72	11·82	11·92	12·03	—	—	—	—	—	—	—	—	—
5	(a)	13·28	13·81	14·30	14·71	15·03	15·25	15·34	15·28	15·06	14·66	14·07	—	—	—	—	—	—	—
	(b)	13·28	13·33	13·39	13·45	13·53	13·61	13·69	13·78	13·87	13·97	14·07	—	—	—	—	—	—	—
6	(a)	15·09	15·58	16·03	16·44	16·79	17·07	17·27	17·37	17·37	17·24	16·98	16·03	—	—	—	—	—	—
	(b)	15·09	15·14	15·20	15·26	15·33	15·41	15·48	15·57	15·65	15·74	15·84	16·03	—	—	—	—	—	—
7	(a)	16·81	17·27	17·70	18·10	18·46	18·77	19·02	19·21	19·32	19·35	19·28	18·84	17·92	—	—	—	—	—
	(b)	16·81	16·87	16·93	16·99	17·06	17·14	17·21	17·29	17·37	17·46	17·55	17·73	17·92	—	—	—	—	—
8	(a)	18·48	18·91	19·32	19·71	20·07	20·39	20·67	20·90	21·08	21·20	21·25	21·13	20·64	19·76	—	—	—	—
	(b)	18·48	18·54	18·60	18·67	18·74	18·81	18·88	18·96	19·04	19·13	19·21	19·39	19·57	19·76	—	—	—	—
9	(a)	20·09	20·50	20·90	21·28	21·64	21·97	22·26	22·52	22·74	22·91	23·03	23·10	22·91	22·41	21·56	—	—	—
	(b)	20·09	20·15	20·22	20·29	20·36	20·44	20·51	20·59	20·67	20·75	20·84	21·01	21·19	21·37	21·56	—	—	—
10	(a)	21·67	22·06	22·45	22·82	23·17	23·50	23·80	24·08	24·32	24·52	24·69	24·90	24·90	24·66	24·15	23·33	—	—
	(b)	21·67	21·73	21·80	21·88	21·95	22·02	22·10	22·18	22·26	22·34	22·42	22·60	22·77	22·95	23·14	23·33	—	—
11	(a)	23·21	23·59	23·97	24·33	24·67	25·00	25·31	25·59	25·85	26·08	26·28	26·57	26·70	26·65	26·38	25·86	—	—
	(b)	23·21	23·28	23·35	23·43	23·50	23·58	23·66	23·74	23·82	23·90	23·98	24·15	24·33	24·51	24·69	24·88	—	—
12	(a)	24·72	25·10	25·46	25·81	26·15	26·48	26·79	27·08	27·35	27·59	27·81	28·16	28·39	28·46	28·37	28·07	26·79	—
	(b)	24·72	24·80	24·87	24·95	25·03	25·11	25·18	25·27	25·35	25·43	25·51	25·68	25·86	26·04	26·22	26·41	26·79	—
13	(a)	26·22	26·58	26·93	27·28	27·62	27·94	28·25	28·54	28·81	29·07	29·30	29·70	29·99	30·16	30·19	30·06	29·22	—
	(b)	26·22	26·29	26·37	26·45	26·53	26·61	26·69	26·77	26·85	26·94	27·02	27·19	27·37	27·55	27·73	27·91	28·29	—
14	(a)	27·69	28·04	28·39	28·73	29·06	29·38	29·69	29·98	30·26	30·52	30·77	31·19	31·53	31·77	31·89	31·88	31·39	30·16
	(b)	27·69	27·77	27·85	27·93	28·01	28·09	28·17	28·25	28·34	28·42	28·51	28·68	28·86	29·03	29·22	29·40	29·77	30·16
15	(a)	29·14	29·49	29·83	30·16	30·49	30·80	30·11	31·40	31·68	31·95	32·20	32·66	33·03	33·32	33·51	33·59	33·37	32·52
	(b)	29·14	29·22	29·30	29·38	29·47	29·55	29·63	29·72	29·80	29·89	29·97	30·15	30·32	30·50	30·69	30·87	31·24	31·62

$M = \mathrm{N} \log_e \left\{ \sum_{t=1}^{k} (\nu_t s_t^2)/\mathrm{N} \right\} - \sum_{t=1}^{k} (\nu_t \log_e s_t^2)$ $\mathrm{N} = \sum_{t=1}^{k} \nu_t$ $c_1 = \sum_{t=1}^{k} \frac{1}{\nu_t} - \frac{1}{\mathrm{N}}$ N.B. $\log_e x = 2{\cdot}3026 \log_{10} x$.

Table 3. *Table of* $C = c_3(a) = c_1^3/k^2$ *and* $\Delta C = c_3(b) - c_3(a) = c_1 - c_1^3/k^2$ *to facilitate interpolation in the tables of percentage points of M*

k \ c_1		0·5	1·0	1·5	2·0	2·5	3·0	3·5	4·0	4·5	5·0	6·0	7·0	8·0	10·0	12·0	14·0
3	C	0·014	0·111	0·375	0·889	1·736	3·000	—	—	—	—	—	—	—	—	—	—
	ΔC	0·486	0·889	1·125	1·111	0·764	0·000	—	—	—	—	—	—	—	—	—	—
4	C	0·008	0·062	0·211	0·500	0·977	1·688	2·680	4·000	—	—	—	—	—	—	—	—
	ΔC	0·492	0·938	1·289	1·500	1·523	1·312	0·820	0·000	—	—	—	—	—	—	—	—
5	C	0·005	0·040	0·135	0·320	0·625	1·080	1·715	2·560	3·645	5·000	—	—	—	—	—	—
	ΔC	0·495	0·960	1·365	1·680	1·875	1·920	1·785	1·440	0·855	0·000	—	—	—	—	—	—
6	C	0·003	0·028	0·094	0·222	0·434	0·750	1·191	1·778	2·531	3·472	6·000	—	—	—	—	—
	ΔC	0·497	0·972	1·406	1·778	2·066	2·250	2·309	2·222	1·969	1·528	0·000	—	—	—	—	—
7	C	0·003	0·020	0·069	0·163	0·319	0·551	0·875	1·306	1·860	2·551	4·408	7·000	—	—	—	—
	ΔC	0·497	0·980	1·431	1·837	2·181	2·449	2·625	2·694	2·640	2·449	1·592	0·000	—	—	—	—
8	C	0·002	0·016	0·053	0·125	0·244	0·422	0·670	1·000	1·424	1·953	3·375	5·359	8·000	—	—	—
	ΔC	0·498	0·984	1·447	1·875	2·256	2·578	2·830	3·000	3·076	3·047	2·625	1·641	0·000	—	—	—
9	C	0·002	0·012	0·042	0·099	0·193	0·333	0·529	0·790	1·125	1·543	2·667	4·235	6·321	—	—	—
	ΔC	0·498	0·988	1·458	1·901	2·307	2·667	2·971	3·210	3·375	3·457	3·333	2·765	1·679	—	—	—
10	C	0·001	0·010	0·034	0·080	0·156	0·270	0·429	0·640	0·911	1·250	2·160	3·430	5·120	10·000	—	—
	ΔC	0·499	0·990	1·466	1·920	2·344	2·730	3·071	3·360	3·589	3·750	3·840	3·570	2·880	0·000	—	—
11	C	0·001	0·008	0·028	0·066	0·129	0·223	0·354	0·529	0·753	1·033	1·785	2·835	4·231	8·264	—	—
	ΔC	0·499	0·992	1·472	1·934	2·371	2·777	3·146	3·471	3·747	3·967	4·215	4·165	3·769	1·736	—	—
12	C	0·001	0·007	0·023	0·056	0·109	0·188	0·298	0·444	0·633	0·868	1·500	2·382	3·556	6·944	12·000	—
	ΔC	0·499	0·993	1·477	1·944	2·391	2·812	3·202	3·556	3·867	4·132	4·500	4·618	4·444	3·056	0·000	—
13	C	0·001	0·006	0·020	0·047	0·092	0·160	0·254	0·379	0·539	0·740	1·278	2·030	3·030	5·917	10·225	—
	ΔC	0·499	0·994	1·480	1·953	2·408	2·840	3·246	3·621	3·961	4·260	4·722	4·970	4·970	4·083	1·775	—
14	C	0·001	0·005	0·017	0·041	[0·080	0·138	0·219	0·327	0·465	0·638	1·102	1·750	2·612	5·102	8·816	14·000
	ΔC	0·499	0·995	1·483	1·959	2·420	2·862	3·281	3·673	4·035	4·362	4·898	5·250	5·388	4·898	3·184	0·000
15	C	0·001	0·004	0·015	0·036	0·069	0·120	0·191	0·284	0·405	0·556	0·960	1·524	2·276	4·444	7·680	12·196
	ΔC	0·499	0·996	1·485	1·964	2·431	2·880	3·309	3·716	4·095	4·444	5·040	5·476	5·724	5·556	4·320	1·804

THE DESIGN OF OPTIMUM MULTIFACTORIAL EXPERIMENTS

By R. L. PLACKETT and J. P. BURMAN

1. Introduction

A problem which often occurs in the design of an experiment in physical or industrial research is that of determining suitable tolerances for the components of a certain assembly; more generally of ascertaining the effect of quantitative or qualitative alterations in the various components upon some measured characteristic of the complete assembly. It is sometimes possible to calculate what this effect should be; but it is to the more general case when this is not so that the methods given below apply. In such a case it might appear to be best to vary the components independently and study separately the effect of each in turn. Such a procedure, however, is wasteful either of labour or accuracy, while to carry out a complete factorial experiment (i.e. to make up assemblies of all possible combinations of the n components) would require L^n assemblies, where L is the number of values (assumed constant) at which each component can appear. For L equal to 2 this number is large for moderate n and quite impracticable for n greater than, say, 10. For larger L the situation is even worse. What is required is a selection of N assemblies from the complete factorial design which will enable the component effects to be estimated with the same accuracy as if attention had been concentrated on varying a single component throughout the N assemblies. Designs are given below for $L = 2$ and all possible $N \leqslant 100$ except $N = 92$ (as yet not known), and for $L = 3, 4, 5, 7$ when $N = L^r$ (for all r).

The following results have been obtained:

(*a*) When each component appears at L values, all main effects may be determined with the maximum precision possible using N assemblies, if, and only if, L^2 divides N, and certain further conditions are satisfied.

(*b*) For $L = 2$, the solution of the problem is for practical purposes complete. In designs of the form $N = L^r$, the effects of certain interactions between the components may also be estimated with maximum precision.

The precision naturally increases with the number of assemblies measured, and to this extent depends on the judgement of the experimenter. Before explaining the procedure in detail, some introductory remarks are necessary on the assumptions made and the method of least squares.

2. Experimental effects when $L = 2$

Each component in the assembly appears at two values throughout; it will be convenient to call one of them the nominal and the other the extreme, where the former usually refers to the actual nominal value and the latter to an extreme of the tolerance range for the component in question (the same extreme for each appearance of the component in a given experiment). Denote the measurable characteristics of the components in the assembly (one per component) by $x_1, x_2, \ldots, x_n$ and the measured assembly characteristic by y.

Then
$$y = y(x_1, x_2, \ldots, x_n),$$
where the functional relationship is in general unknown. Suppose that the nominal value of x_i is x_i^0 and of y, y_1. Thus
$$y_1 = y(x_1^0, x_2^0, \ldots, x_n^0).$$

Suppose also that the extreme value of x_i under consideration is x'_i. Then the *main effect* of component 1 is $m_1 = [\Sigma y(x'_1, x_2, x_3, \ldots, x_n) - \Sigma y(x^0_1, x_2, x_3, \ldots, x_n)]/2^n$,

where the total number of possible assemblies is 2^n. In each of the two summations above, the indices on the x_i $(i \neq 1)$ range over all possible sets of values. Similarly, $m_2, m_3, \ldots m_n$ are defined. For brevity the above equation will be written:

$$m_1 = [\Sigma y(x'_1) - \Sigma y(x^0_1)]/2^n,$$

and in general $\Sigma y\ (x'_i x'_j \ldots x^0_k)$ will represent the function y evaluated with $x_i x_j \ldots x_k$, taking the values shown and summed over all possible sets of values of the variables that have been suppressed. The main effect of a component is thus seen to be the mean effect on the measured assembly characteristic which that component would produce if acting on its own. Proceeding further we define the *interaction* between components 1, 2, 3, ..., p as

$$m_{(123\ldots p)} = [\Sigma y(x'_1 x'_2 \ldots x'_p) - \Sigma\Sigma y(x'_1 x'_2 \ldots x'_{p-1} x^0_p) + \Sigma\Sigma y(x'_1 x'_2 \ldots x'_{p-2} x^0_{p-1} x^0_p) + \ldots + (-1)^p\, \Sigma y(x^0_1 x^0_2 \ldots x^0_p)]/2^n,$$

where the inner summation is as explained above; the outer extends over the ${}_pC_1$, ${}_pC_2$, etc., selections of 1, 2, 3, etc., indices 0 available. The nature of an interaction has been discussed by Fisher (1942) and others, and our definition accords with the usual one.

If main effects are regarded as being of the first order of small quantities and *if the function y may be differentiated*, the first approximation to $m_{(123\ldots p)}$ is

$$m_{(123\ldots p)} = (\partial^p y/\partial x_1 \partial x_2 \ldots \partial x_p)\,(x'_1 - x^0_1)\,(x'_2 - x^0_2) \ldots (x'_p - x^0_p),$$

the derivative being *averaged over the values it takes for all sets of values* of the remaining components. This shows that when the variables are measured on a continuous scale we may validly neglect all the interactions above a certain order, for a $(p-1)$th order interaction (one in p components) is of the pth order of smallness. But the justification for this assumption when some of the x_i are qualitative and not quantitative (and it is frequently made) must be found in considerations outside the data which the experiment provides, in commonsense or philosophical grounds.

The grand mean $M = \Sigma y(x_1, x_2, \ldots, x_n)/2^n$ where the summation is over all possible sets of values of the components. In the jth assembly of an actual experiment, some components will be at nominal and some at their extreme values. If the true value of the assembly characteristic is then y_j, it is found on solving the above equations that

$$y_j = M + a_{j1} m_1 + a_{j2} m_2 + \ldots + a_{jn} m_n + a_{j,\,n+1} m_{(12)} + \ldots + a_{j,2^n} m_{(123\ldots n)}, \tag{1}$$

where the coefficient of m_i is ± 1 according as the ith component is at extreme or nominal in the jth assembly; the coefficient of $m_{(123\ldots p)}$ is ± 1 according as the number of plus ones among the coefficients of $m_1 m_2 \ldots m_p$ is odd or even. In doing this the signs of the odd-order interactions (involving an even number of factors) have been reversed, but the notation is convenient, for then the coefficients in y_1 are all minus one. It is assumed that y_1 is always one of the selected assemblies, and this is no real restriction upon the design.

3. Least squares and precision

The purpose of the experiment is to estimate those of the quantities m as may not be assumed negligible from a set of measurements $r_1, r_2, \ldots, r_N$. For this we must solve a set of N linear equations represented by (1). The equations always involve M, and therefore to estimate q of the m's it is necessary to make at least $(q+1)$ measurements. If exactly $(q+1)$ assemblies

are measured, there is a unique set of m's satisfying the equations; if more than $(q+1)$ are measured there will be no unique solution and the best estimates are, as is well known, obtained by the method of least squares. This obtains the set of m's which minimizes

$$S = \sum_j (r_j - M - a_{j1}m_1 - a_{j2}m_2 \ldots - a_{jn}m_n \ldots)^2,$$

where r_j is the measurement in the jth assembly whose true value is y_j. Normally $(q+1)$ is much less than 2^n, all high-order interactions being neglected, so that the number of assemblies N may be made much smaller than for the complete factorial design.

As already stated, the greater the number of assemblies measured, the greater the precision with which component effects may be estimated. On account of errors of measurement and the neglect of certain effects the minimum S_0 of S is not zero. In fact $S_0/(N-q-1)$ provides an unbiased estimate s^2 of σ^2, the variance of error of each measurement (assumed the same for all assemblies). The error variance in the estimation of an effect m_i is of the form σ^2/t_i, where t_i is called the precision constant. It depends only on the design of the experiment, and can be increased indefinitely by increasing N. Our object is to find designs which maximize all the t_i simultaneously for given N. They will be called optimum designs. The ratio of m_i to $s/\sqrt{t_i}$ has a t-distribution on the null hypothesis—that the true value of m_i is zero. The effect of increasing the precision is, first, to increase the power of the t-test in detecting any departure of m_i from zero; secondly, to increase the accuracy of its estimation. In the designs given at the end of this paper, for $L = 2$ all main effects may be estimated with maximum precision N (given N assemblies), that is, the standard error of $m_i = \sigma/\sqrt{N}$ provided N is a multiple of 4. In cases where $N = 2^r$ certain interactions may also be estimated with the maximum precision. The choice of N (subject to $N \geqslant q+1$) will depend on the extent to which the experimenter wishes to minimize the effect of his experimental error.

4. Requirements for optimum designs (any L)

I. Consider now the case of n components each of which may take L values. If interactions are neglected, the true values y_i may be assumed linear functions of certain constants representing the main effects, as was proved rigorously for the case $L = 2$. In general let $x_{j(l)}$ represent the effect due to the jth component at its lth value. The true value of the measurement on the ith assembly is

$$y_i = \sum_j x_{j(l)} \begin{bmatrix} i = 1, 2, \ldots, N \\ j = 1, 2, \ldots, n \\ l = 1, 2, \ldots, L \end{bmatrix},$$

where l represents the value at which the jth component appears in the ith assembly. We now introduce certain new variables in terms of which to express the $x_{j(l)}$, as the primary interest is in the *change* of assembly characteristic caused by certain *changes* in the components.

Let Q be a non-singular $L \times L$ matrix whose first column consists entirely of ones, such that $Q = OD$, where O is orthogonal and D diagonal. The condition on the first column of Q implies that $d_{11} = \sqrt{L}$.

Let
$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_L \end{bmatrix} = Q^{-1} \begin{bmatrix} x_{1(1)} \\ x_{1(2)} \\ \vdots \\ x_{1(L)} \end{bmatrix} = Q^{-1}X_1,$$

$u_1 = (x_{1(1)}+x_{1(2)}+\ldots+x_{1(L)})/L$ = mean effect of component 1 and $u_2, u_3, \ldots, u_L$ are constants which determine the effect of *changes* in this component upon the assembly characteristic. The orthogonality property will be used later. Therefore

$$X_1 = QU = \begin{bmatrix} u_1+a_{12}u_2+a_{13}u_3+\ldots+a_{1L}u_L \\ u_1+a_{22}u_2+a_{23}u_3+\ldots+a_{2L}u_L \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ u_1+a_{L2}u_2+a_{L3}u_3+\ldots+a_{LL}u_L \end{bmatrix},$$

where the a_{ij} are certain constants. Similarly, introduce variables $v_1, v_2, v_3, \ldots, v_L$ for component 2 and write:

$$X_2 = \begin{bmatrix} x_{2(1)} \\ x_{2(2)} \\ \ldots \\ x_{2(L)} \end{bmatrix} = \begin{bmatrix} v_1+a_{12}v_2+a_{13}v_3+\ldots+a_{1L}v_L \\ v_1+a_{22}v_2+a_{23}v_3+\ldots+a_{2L}v_L \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ v_1+a_{L2}v_2+a_{L3}v_3+\ldots+a_{LL}v_L \end{bmatrix},$$

where v_1 is the mean effect of component 2. And so on. Hence

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \ldots \\ y_N \end{bmatrix} = AX, \quad \text{where} \quad X = \begin{bmatrix} M \\ u_2 \\ \ldots \\ u_L \\ v_2 \\ \ldots \\ v_L \\ \ldots \end{bmatrix},$$

$$M = u_1+v_1+\ldots \text{ to } n \text{ terms.}$$

A is a matrix with N rows and $n(L-1)+1$ columns, the first column consisting of ones, and the remainder consisting of the elements a_{ij} belonging to Q. The columns fall into sets (corresponding to the components) of $(L-1)$ after the first, and the rows of the submatrix formed by such a set consist of repetitions of the rows of Q. At this stage renumber the suffices of the a_{ij} so that A may be written (a_{ij}).

II. The vector Y is known. Solving the equations by least squares (assuming $N \geqslant n(L-1)+1$) gives the so-called normal equations $A'Y = A'AX = CX$ say, i.e. $X = C^{-1}A'Y$. If σ^2 is the error variance of a single observation y, it is proved in text-books that

$$\operatorname{var}(\theta_k) = |C_{kk}|\sigma^2/|C|,$$

where θ_k is the kth element of X and C_{kk} the cofactor of c_{kk} in $C = [c_{ij}]$, C being a symmetric $n \times n$ matrix. It is required to minimize $|C_{kk}|/|C|$, i.e. to maximize $t = |C|/|C_{kk}|$ by suitable choice of design.

Write $c_{ij}/c_{ii}^{\frac{1}{2}}c_{jj}^{\frac{1}{2}} = r_{ij}$ and the matrix $R = [r_{ij}]$, where $r_{ii} = 1$ and $r_{ij} = r_{ji}$.

Now
$$r_{ij} = \sum_r a_{ri}a_{rj} \Big/ \left(\sum_r a_{ri}^2\right)^{\frac{1}{2}} \left(\sum_r a_{rj}^2\right)^{\frac{1}{2}}.$$

If $(a_{1i}, a_{2i}, \ldots, a_{Ni})$ and $(a_{1j}, a_{2j}, \ldots, a_{Nj})$ be interpreted as the co-ordinates of two points P_i and P_j in a Euclidean N-space, then $r_{ij} = \cos P_iOP_j$, where O is the origin and hence $r_{ij}^2 \leqslant 1$.

Now $t = |R|S_k/|R_{kk}|$, where $S_k = \sum_r a_{rk}^2$ and so S_k must be fixed otherwise t may be increased indefinitely. This is equivalent to fixing the scale of measurement, the preceding section having dealt with the choice of origin at the mean. Eliminate the pth row and column from $|R|$ and $|R_{kk}|$ by pivotal condensation: multiply the pth column by r_{pj} $(p \neq k)$ and subtract from the jth column for all $j \neq p$. A row of zeros appears in the pth row except in the diagonal place where there is a one. The determinants have been reduced in order by one, and the second is still a principal minor of the first.

Thus
$$|r_{ij}| = |r_{ij} - r_{ip}r_{jp}| \quad (\text{remembering } r_{pj} = r_{jp})$$
$$= |(1-r_{ip}^2)^{\frac{1}{2}}(1-r_{jp}^2)^{\frac{1}{2}}r_{ij.p}|$$
defining $r_{ij.p}$ in this manner, where the suffices appearing after the dot represent columns that have been eliminated.

Therefore taking out factors from rows and columns
$$t = S_k \prod_{i \neq p}(1-r_{ip}^2)\left|r_{ij.p}\right| \Big/ \prod_{i \neq p,k}(1-r_{ip}^2)\left|[r_{ij.p}]_{kk}\right|$$
$$= S_k(1-r_{kp}^2)\,|r_{ij.p}|/|[r_{ij.p}]_{kk}|.$$

Now
$$r_{ij.p} = (\cos P_iOP_j - \cos P_iOP_p \cos P_jOP_p)/\sin P_iOP_p \sin P_jOP_p,$$
which is the formula for the cosine of the projection of angle P_iOP_j on to the $(N-1)$-space orthogonal to OP_p. Therefore
$$r_{ij.p}^2 \leqslant 1 \quad (i \neq j) \quad \text{and} \quad r_{ii.p} = 1.$$
The method has obtained a ratio of two determinants of the same type as before, and the process is repeated, step by step, until that in the numerator is of the form
$$\begin{vmatrix} 1 & r_{qk.p_1p_2} & \dots \\ r_{qk.p_1p_2} & \dots & 1 \end{vmatrix}$$
and the denominator is 1. Row and column p_1, p_2, ..., are eliminated in turn (no p being equal to k), and so
$$t = S_k(1-r_{kp_1}^2)(1-r_{kp_2.p_1}^2)(1-r_{kp_3.p_1p_2}^2)\dots(1-r_{kq.p_1p_2\dots p_{n-2}}^2).$$
This is a maximum only when $r_{kp} = 0$ for all $p \neq k$ and all k. For equal precision S_k must be constant for all k and $t = S_k$. Therefore $A'A = C = tI$. Hence the designs for which the maximum precision is attained are those which correspond to columns of an orthogonal matrix (apart from an arbitrary multiplier).

At this point it is convenient to prove the formula for the error variance. Let A be the non-square matrix with orthogonal columns of the equations: $Y = AX$. Introduce further columns U so that (A, U) is a square orthogonal matrix, and corresponding dummy variables whose column vector is Z. The least squares solution of the above equations is X_0 given by $A'AX_0 = A'Y$, therefore
$$tIX_0 = A'Y, \quad X_0 = \frac{1}{t}A'Y. \tag{1}$$
The equations $[A, U]\begin{bmatrix} X \\ Z \end{bmatrix} = Y$ have a unique solution, and on multiplying by $\begin{bmatrix} A' \\ U' \end{bmatrix}$,
$$tI\begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} A'Y \\ U'Y \end{bmatrix},$$
so the resulting value of $X = X_0$ as before. The residual vector $E = Y - AX_0 = UZ$.
$$\text{Sum of squares of residuals} = E'E = Z'U'UZ = tZ'Z = t\Sigma z_i^2. \tag{2}$$

III. Consider now any pair of components f and g. Suppose they appear together at their lth and l'th values respectively $w_{ll'}$ times. This defines an $L \times L$ matrix $W = [w_{ll'}]$. The scalar product of a column of A belonging to f by a column of A belonging to g is zero by the orthogonality of A. Let these correspond to the uth and vth columns of Q and revert to the old suffices of a_{ij} corresponding to Q, i.e. $Q = [a_{ij}]$, where $i, j = 1, 2, \ldots, L$.

Then $a_{lu} . w_{ll'} . a_{l'v}$ in dummy suffices equals

$$Q'WQ = \begin{bmatrix} N & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & 0 \end{bmatrix}.$$

The N appears because the first column for f is the same as the first column for g, equal to the first column of A which consists entirely of ones.

Now $Q = OD$ where O is orthogonal and D diagonal, therefore $Q'WQ = DO'WOD$, i.e.

$$O'WO = D^{-1} \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} D^{-1}$$

$$= \begin{bmatrix} N/d_{11}^2 & 0 \\ 0 & 0 \end{bmatrix}$$

Therefore

$$W = OO'WOO' = \begin{bmatrix} 1/\sqrt{L} & \\ 1/\sqrt{L} & \text{other terms} \\ 1/\sqrt{L} & \end{bmatrix} \begin{bmatrix} N/L & 0 \\ & \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{L} & 1/\sqrt{L} & \ldots & 1/\sqrt{L} \\ & \text{other terms} & & \\ & & & \end{bmatrix}$$

$$= \begin{bmatrix} N/L^2 & N/L^2 & \ldots & N/L^2 \\ N/L^2 & N/L^2 & \ldots & N/L^2 \\ \ldots & \ldots & \ldots & \ldots \\ N/L^2 & N/L^2 & & N/L^2 \end{bmatrix}.$$

Sum of terms in lth row is the number of replications of the lth value of f. Therefore

(i) Each component is replicated at each of its values the same number of times.

(ii) Each pair of components occur together at every combination of values the same number of times.

(iii) The number of assemblies is divisible by the square of the number of values.

The converse—that under these conditions the matrix A is orthogonal—can be proved by reversing these steps. The actual matrix Q chosen is unimportant, and the design can be specified by means of a rectangular array with N rows and n columns containing L different letters $(a, b, c, \ldots, k)$ representing the L values of each component. The problem is then a purely combinatorial one. If $N = KL^2$, the maximum number of columns n is

$$(KL^2 - 1)/(L-1)$$

or its integral part since $KL^2 \geqslant n(L-1)+1$. We propose to call designs of this type *multifactorial designs*.

Returning to the case of $L = 2$, it is necessary to obtain an orthogonal $4K \times 4K$ matrix A whose first column consists entirely of ones. Choosing $Q = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ the other columns of A consist of equal numbers of $+1$ and -1. The signs may be changed down the length of certain columns without spoiling the design so that the first row apart from the corner element consists entirely of -1. Then apart from this row (in future called the basic row) and the first column the design consists of a square matrix with $2K$ plus and $(2K-1)$ minus ones in each column and (by orthogonality) row, and such that each pair of columns contains a pair of plus ones in the same row K times. The estimates of component effects are obtained from equation (1) of § 4 (II):

$$X_0 = \frac{1}{4K} A'Y \quad (\text{here } t = 4K).$$

Thus they can be evaluated by addition and subtraction with only one division. This simplicity appears in the illustrative example given in §§ 9 and 10. The dummy variables z_i are similarly evaluated and the estimated error variance is

$$s^2 = \frac{4K}{4K-n-1} \Sigma z_i^2.$$

5. Methods of solution

Certain methods of constructing orthogonal matrices with elements plus or minus one are known (Paley, 1933). They depend upon the theory of finite fields, an outline of which will now be given.

A field F is defined as a set of quantities which is closed with respect to two operations, addition and multiplication (i.e. if a, b in F, so are $a+b, ab$). These quantities satisfy the following laws:

$$\text{(i)}\ a+b = b+a. \qquad \text{(ii)}\ a+(b+c) = (a+b)+c. \qquad \text{(iii)}\ a(b+c) = ab+ac.$$
$$ab = ba. \qquad a(bc) = (ab)c.$$

(iv) There is an x such that $a+x = b$ for every a, b.

From these it may be proved: (a) There is a unique quantity 0 such that $a+0 = a$ for all a. (b) The quantity x in (iv) is unique. (c) $a.0 = 0$. Finally, we add

(v) There is a y such that $ay = b$ for every b, all $a \neq 0$, to our axioms.

Hence as before: (d) There is a unique quantity 1 such that $a.1 = a$ for all a. (e) There is a unique quantity a^{-1} such that $a.a^{-1} = 1$ $(a \neq 0)$. (f) The quantity y in (v) is unique. $y = a^{-1}b$.

Consider the integers $0, 1, 2, \ldots, (p-1)$ where p is prime, and write $a = b$ if $(a-b)$ is divisible by p. Then this set of integers forms a finite field as may be easily shown. For example, when $p = 5$, the numbers in the field are 0, 1, 2, 3, 4.

$$2+4 = 6 = 1, \quad 2+3 = 5 = 0, \quad 2.3 = 4.4 = 1.$$

Hence 2 and 3 are reciprocals and 4 is its own reciprocal. This field is called the Galois field of order p, $GF(p)$.

Now suppose x is a number algebraic over $GF(p)$, that is, x satisfies an algebraic equation with coefficients in $GF(p)$. Then it defines an algebraic extension of $GF(p)$, namely, all polynomials in x with coefficients in $GF(p)$. If x satisfies an equation irreducible in $GF(p)$ and of degree n, there are p^n distinct polynomials in x. They are of the form

$$f(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \quad (a_0, a_1, \ldots, a_{n-1} \text{ in } GF(p)).$$

Such an algebraic extension is, in fact, a field. Moreover, all fields of degree n over $GF(p)$ may be shown to be equivalent. There is a member of the extended field α such that $1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{q-2}$ $(q = p^n)$ constitute the non-zero elements of the field and $\alpha^{q-1} = 1$; an extension of *Fermat's theorem*. Any one of these equivalent fields is denoted by $GF(p^n)$. We shall now require various simple theorems.

DEF. If a non-zero element of a finite field is a perfect square $(a = b^2)$ it is called a quadratic residue of the field. All other non-zero elements are non-Q.R.'s.

TH. 1. The numbers of Q.R.'s and non-Q.R.'s are equal $(p > 2)$.

For every $b = \alpha^u$, $\qquad a = b^2 = \alpha^{2u} = \alpha^{2u-\lambda(q-1)} \quad (\lambda \text{ integral}).$

But $(q-1)$ is even $(p > 2)$. Hence only even powers of α are Q.R.'s.

Therefore there are $\frac{1}{2}(q-1)$ Q.R.'s. and $\frac{1}{2}(q-1)$ non-Q.R.'s.

We now define the Legendre function $\chi(a)$:

$$\begin{aligned} \chi(0) &= 0, \\ \chi(a) &= +1 \quad \text{when } a \text{ is a Q.R.} \\ &= -1 \quad \text{when } a \text{ is a non-Q.R.} \end{aligned}$$

Th. 1 states that $\sum_a \chi(a) = 0$ (summation over whole field).

TH. 2. $\chi(a)\,\chi(b) = \chi(ab)$.

This is trivial when $a = 0$ or $b = 0$. Otherwise $a = \alpha^u, b = \alpha^v$ and $ab = \alpha^{u+v}$ is a Q.R. if and only if $(u+v)$ is even, i.e. u and v of same parity. This proves the result.

TH. 3. $\left.\begin{aligned} \chi(-1) &= +1 \text{ if } q = 4t+1 \\ &= -1 \text{ if } q = 4t-1 \end{aligned}\right\}$ for integral t.

For $\alpha^{q-1} = +1$.

Therefore $\alpha^{\frac{1}{2}(q-1)} = \pm 1 = -1$ since powers of α are distinct up to α^{q-1}.

Hence -1 is a Q.R. if and only if $\frac{1}{2}(q-1)$ is even $= 2t$ and $q = 4t+1$.

TH. 4. $\sum_j \chi(j-i_1)\chi(j-i_2) = -1$ (summation over all j in $GF(p^n)$; $p > 2$; $i_1 \neq i_2$),

$$\sum_j \chi(j-i_1)\chi(j-i_2) = \sum_j \chi\{(j-i_1)(j-i_2)\} \quad \text{by Th. 2.}$$

Put $$u = j - \frac{(i_1+i_2)}{2}, \quad u_0 = \frac{i_1-i_2}{2} \neq 0.$$

Expression $= \sum_u \chi(u^2 - u_0^2)$ (j is summed over whole field so u will be also).

Put $u = u_0 v \quad (u_0 \neq 0)$.

Expression $= \sum_v \chi\{u_0^2(v^2-1)\}$ (u is summed over whole field so v will be also)

$$= \sum_v \chi(u_0^2)\,\chi(v^2-1) = \sum_v \chi(v^2-1) \quad \text{by Th. 2.}$$

Now if $v^2 - 1 = x^2$, $v^2 - x^2 = 1$,

$$(v-x)(v+x) = 1.$$

If $v+x = y$, $v-x = y^{-1}$.

Therefore $$v = \tfrac{1}{2}(y+y^{-1}), \quad x = \tfrac{1}{2}(y-y^{-1}).$$

Hence the number of values of v for which $\chi(v^2-1) = +1$ or 0 is the number of values of v for which $v = \frac{1}{2}(y+y^{-1})$.

Now if $\quad y+y^{-1} = w+w^{-1}, \quad y^2w+w = yw^2+y, \quad (y-w)(1-yw) = 0.$

Therefore $w = y$ or y^{-1}. y and y^{-1} are distinct unless $y = \pm 1$, $v = \pm 1$ when

$$\chi(v^2-1) = \chi(0) = 0.$$

Hence to every one of the $\frac{1}{2}(q-1)$ reciprocal pairs (y, y^{-1}) corresponds a distinct value of v.
Thus there are $\frac{1}{2}(q-3)$ values of v for which $\chi(v^2-1) = +1$ (excluding $v = \pm 1$).
There are two values of v for which $\chi(v^2-1) = 0$ $(v = \pm 1)$.
Hence there are $\frac{1}{2}(q-1)$ values of v for which $\chi(v^2-1) = -1$.

Therefore
$$\sum_j \chi(j-i_1)\,\chi(j-i_2) = \sum_v \chi(v^2-1) = -1.$$

Applications

I. Consider the matrix $A = (a_{ij})$ $(i, j = 0, 1, 2, \ldots, p)$ of order $(p+1)$, where $p = 4t-1$.

$$a_{i0} = a_{0j} = +1,$$
$$a_{ij} = \chi(j-i) \quad (i \neq 0, j \neq 0, i \neq j),$$
$$a_{ii} = -1.$$

The scalar product of 1st and $(i+1)$th rows

$$= a_{00}a_{i0} + a_{0i}a_{ii} + \sum_{j=1}^{p} \chi(j-i)$$
$$= 1 - 1 + 0 = 0 \quad \text{(Th. 1)}.$$

Scalar product of (i_1+1)th and (i_2+1)th rows

$$= a_{i_10}a_{i_20} + a_{i_1i_1}a_{i_2i_1} + a_{i_2i_2}a_{i_1i_2} + \sum_{j=1}^{p} \chi(j-i_1)\,\chi(j-i_2)$$
$$= 1 - \chi(i_1-i_2) - \chi(i_2-i_1) - 1 \quad \text{(Th. 4)}$$
$$= 0 \text{ since } p = 4t-1 \quad \text{(Th. 3)}.$$

Hence the matrix A is orthogonal.

II. To construct A of order $p^n+1 = 4t$, we associate the rows and columns (except the first) with the elements of $GF(p^n)$ and the proof runs exactly as before.

III. If A is orthogonal $\begin{bmatrix} A & A \\ A & -A \end{bmatrix}$ is also orthogonal and has double the order of A.

Hence an orthogonal matrix A of order $2^h(p^n+1)$ (where $p^n = 4t-1$) or 2^h can be constructed by successive doubling.

IV. If $p^n = 4t+1$, (p^n+1) is not divisible by 4.

But an A of order $2(p^n+1)$ can be obtained by a slight modification of the method.

Consider the matrix $B = (b_{ij})$ $(i, j = 0, 1, 2, \ldots, p^n)$ of order (p^n+1) $[p^n = 4t+1]$.

$$b_{i0} = b_{0j} = +1 \qquad (i \neq 0, j \neq 0),$$
$$b_{ij} = \chi(u_j - u_i) \quad (i \neq 0, j \neq 0)$$

where u_i is the element of $GF(p^n)$ associated with the $(i+1)$th row and column of B, $i = 1, 2, \ldots, p^n$,

$$b_{00} = 0.$$

Scalar product of 1st and $(i+1)$th rows

$$= b_{00}b_{i0} + \sum_{j=1}^{p^n} \chi(u_j - u_i) = 0 \quad \text{(Th. 1)}.$$

Scalar product of (i_1+1)th and (i_2+1)th rows

$$= b_{i_10}b_{i_20}+b_{i_1i_1}b_{i_2i_1}+b_{i_2i_2}b_{i_1i_2}+\sum_{j=1}^{p^n}\chi(u_j-u_{i_1})\,\chi(u_j-u_{i_2})$$

$$= 1-1 = 0 \quad \text{(Th. 4)}.$$

Thus B is orthogonal.

Now replace $+1$ by the submatrix $C = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}$.

-1 by the submatrix $-C = \begin{bmatrix} -1 & -1 \\ -1 & +1 \end{bmatrix}$.

0 by the submatrix $D = \begin{bmatrix} +1 & -1 \\ -1 & -1 \end{bmatrix}$. The new matrix A thus formed is of order $2(p^n+1)$.

Consider the scalar products of the $(2i_1+1)$th and $(2i_1+2)$th rows with the $(2i_2+1)$th and $(2i_2+2)$th rows. This is a (2×2) matrix $M_{i_1i_2}$.

Now
$$M_{i_1i_2} = \sum_{j=0}^{p^n}(b_{i_1j}C)(b_{i_2j}C')+(D)(b_{i_2i_1}C')+(b_{i_1i_2}C)(D') \quad [j\neq i_1, j\neq i_2]$$

$$= CC'\sum_{j=0}^{p^n}b_{i_1j}b_{i_2j}+X(u_{i_1}-u_{i_2})(DC'+CD') \quad [j\neq i_1, j\neq i_2]$$

(by Th. 2 and Th. 3 for $p^n = 4t+1$)

$$= CC'0+\chi(u_{i_1}-u_{i_2})\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

since $\sum_{j=0}^{p^n}b_{i_1j}b_{i_2j} = 0$ (orthogonality of B) and the omitted terms vanish

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, it is clear that the $(2i_1+1)$th and $(2i_1+2)$th rows are orthogonal to each other. *Hence A is orthogonal and of the type required.*

Thus by successive doubling we may obtain matrices of order $2^h(p^n+1)$ where $p^n = 4t+1$.

V. Summing up we have:

If $N = 2^h(p^n+1) = 4K$, where p is an odd prime or zero, an orthogonal matrix A can be constructed with plus and minus ones.

The matrices constructible by these methods include all values of $N = 4K$ up to 100 excepting 92. Those of order 2^r are structurally the same as the complete factorial design in r factors *if they have been obtained by successive doubling*. These will be called geometrical designs because of their close connexion with finite geometries. It is clear that if two columns of the design represent main effects, and if the interaction column corresponding to them in the complete factorial case is a dummy in the actual experiment, it may be used to estimate the interaction. The condition for a column to be the interaction between p other columns is that it is $\pm D$, where D is a column vector whose ith element is the product of the ith elements of the original p columns. So far interaction columns have only been found in the geometrical designs and in them every interaction between an arbitrary set of columns is a column of the design. It must also be mentioned that the cyclic designs for $N = 2^r$ obtainable by the method of § 8 depending on $GF(2^r)$ are in fact merely permutations of the geometric designs. They are the forms used in the tables for convenience.

6. CASE OF MORE THAN TWO LEVELS

We now provide experimental designs for determining component effects with maximum precision when the number of values L is greater than 2. These solutions cover the cases where the number of assemblies $N = L^r$, L being a prime or a power of a prime and r any positive integer. Two methods are given: in the first, successive columns of the design are formed by simple operations on the preceding columns; in the second, which is of more limited application, the design is specified by one column, all others being cyclic permutations of this.

7. MODIFIED FACTORIAL DESIGNS

The methods given in this section for the construction of multifactorial designs, although discovered independently, are nevertheless identical with those used by Bose & Kishen (1940) to express the generalized interaction for the purpose of confounding certain contrasts with block differences in agricultural experiments. They construct their interactions directly from finite projective geometries without using, as we have done, the intermediate device of orthogonal sets of Latin squares. We shall, however, describe these methods, as they may not be familiar to experimenters in this country, especially not in the way in which we propose using them.

Suppose a complete factorial experiment is carried out for r factors each at L levels (i.e. in this case r components measured at L values) so that L^r assemblies are made. Let the levels be called $0, 1, 2, \ldots, (L-1)$. Then the r main effects define r columns of a design array (with L^r rows) containing these L symbols. Each symbol appears the same number of times in a column as any other. Each combination of symbols for two columns occurs equally frequently. We shall apply the term orthogonal to such a pair of columns. Now let A, B be two orthogonal columns. The interaction AB has $(L-1)^2$ degrees of freedom. Since each column of the array is associated with $(L-1)$ degrees of freedom, a first-order interaction is represented by a set of $(L-1)$ columns which will be called the terms of this interaction. Similarly, an interaction of the mth order (i.e. involving $m+1$ factors) is represented by $(L-1)^m$ columns of the design array.

Now an interaction between two factors is most naturally defined by the following conditions:

(*a*) Each combination of levels of A and B corresponds to only one level within each term of AB.

(*b*) The terms of AB are orthogonal to A and B and to one another.

Condition (*a*) means that if, for instance, in one assembly level 2 of A and level 5 of B occur together, and if a term of AB is defined to appear at level 3 in this case, then whenever A and B occur at these levels together again, this term of AB appears at level 3. Since, owing to the complete factorial basis of the design, every combination of levels of ABC occurs equally often, each combination of levels of A and B occurs equally often with every level of C. But such a combination of levels of A and B fixes the level in a term of AB by condition (*a*). Hence each level in a term of AB occurs equally often with every level of C. In other words, each interaction term will be orthogonal to the main effects not connected with it.

Now for condition (*b*). If the rows and columns of a square $L \times L$ array correspond respectively to the L levels of A and B, each cell may be filled up with the level appropriate to a particular interaction term. For any particular term of AB such a square will be Latin, because, regarding a row, the level of A is fixed; all the levels of the interaction term must

occur equally often with this level of A and hence each symbol appears once in every row of the square; similarly it appears once in every column in order that the interaction term may be orthogonal to B. Finally, superimposing the Latin squares for two terms of the same interaction, each symbol belonging to the first term must appear once in the same cell with each symbol of the second term, in order that these two may be orthogonal: thus if conditions (a) and (b) are satisfied the interaction terms are founded upon a completely orthogonal set of Latin squares.

It remains to show that the terms from two different interactions are orthogonal. This follows because every combination of levels of four factors $ABCD$ occurs equally frequently, i.e. each combination of levels of A and B occurs equally often with every combination of levels of C and D. The former correspond to levels in the terms of AB; the latter to levels in terms of CD. Hence a term of AB is orthogonal to a term of CD. Similarly AB and AC may be dealt with. This shows that the first order interaction terms may be joined to the main factors as part of the balanced design. Higher order interactions may be regarded as first order interactions between those of lower order, e.g. $(ABC) = (AB)(C)$, it being understood that $(L-1)$ terms are derived from each term of (AB) taken with the factor C, so that in this case there will be $(L-1)^2$ terms for the second order interaction. This procedure builds up the design by an inductive process, and when the interaction of the $(r-1)$th order has been obtained, it will be complete. The total degrees of freedom in the original factorial design $= L^r - 1$. Hence the number of factors that may be measured if interactions are neglected

$$= \frac{L^r - 1}{L-1}.$$

It may be remarked that there is nothing new in this treatment of the complete factorial design except the modification of the usual Fisher interactions so that they may be placed on the same footing as main effects.

If L is a prime number, cyclic Latin squares exist forming an orthogonal set: each square is obtained by writing the first row of symbols in standard order, successive rows being obtained by shifting the symbols along ρ places from each row to the next $(\rho = 1, 2, \ldots, (L-1))$. In this case the appropriate column of the design is formed as follows: assuming the first row in the order $0, 1, 2, \ldots, (L-1)$, if x level of A and y level of B occur together, the corresponding level for this interaction term is $(y+\rho x)$, the symbols being reduced with modulus L. The squares for $\rho = 1, 2, \ldots, (L-1)$, give all the interaction terms.

When L is the nth power of a prime p, then we associate the L levels of a factor with the elements of a Galois field, $GF(p^n)$. Suppose these elements to be $u_0, u_1, u_2, \ldots, u_{L-1}$ where u_0 is the zero and u_1 the unity of the field. If then u_x level of A and u_y level of B occur together, the corresponding levels for interaction terms are $u_y + u_\rho u_x$, and the squares for $u_\rho = u_1, u_2, \ldots, u_{L-1}$ give all the terms present. The method of constructing completely orthogonal sets of Latin squares from Galois fields is given in Stevens (1939).

For $L = 6$ it is known that no pair of orthogonal Latin squares exists so it is not amenable to this treatment. The design for $N = 9$, $L = 3$ is given below; the accompanying key refers to the column vectors and the rows are labelled as if belonging to a complete factorial design.

	A	B	$(AB)_1$	$(AB)_2$		A	B	$(AB)_1$	$(AB)_2$		A	B	$(AB)_1$	$(AB)_2$
a_1b_1	0	0	0	0	a_2b_1	1	0	1	1	a_3b_1	2	0	2	2
a_1b_2	0	1	1	2	a_2b_2	1	1	2	0	a_3b_2	2	1	0	1
a_1b_3	0	2	2	1	a_2b_3	1	2	0	2	a_3b_3	2	2	1	0

Key: $(AB)_1 = (A) + (B)$, $(AB)_2 = (A) + 2(B)$.

For $r = 4$, $N = 81$, if the main effects are taken as A, B, C, D, $(ABCD)_1$, then all first order interactions are determinable.

8. CYCLIC SOLUTIONS WHEN L IS A PRIME NUMBER

We shall again be concerned in this section with Galois fields: each element of $GF(p^n)$ will be represented by a set of n ordered numbers, each number being $0, 1, 2, \ldots, (p-1)$ where p is prime. Consider the block B_1 of elements which have the integer r in position s $(r = 0, 1, 2, \ldots, p-1;\ s = 1, 2, 3, \ldots, n)$.

For example, in $GF(3^2)$ the block having 2 in position 1 is 20, 21, 22. We require to show that if the elements of this block are multiplied in succession by any element of the field other than $000 \ldots 0t$ $(1 \leqslant t \leqslant p-1)$, then the elements of the resulting block B_2 have equal numbers of all possible r in position s. There are in fact p^{n-1} elements in B_1, and we need to prove that B_2 is subdivisible into

p^{n-2} elements having 0 in position s

p^{n-2} elements having 1 in position s

....

p^{n-2} elements having $p-1$ in position s.

There is no loss in generality if we consider $s = 1$, i.e. we refer now to the first members of all elements of the field. Take now all the elements having 1 in this position. Multiply this block A_1 by any element b of the field and obtain block A_2. Suppose in A_2 that r_0 first members are 0, r_1 first members are $1, \ldots,$ and r_{p-1} first members are $p-1$. Clearly $\sum_0^{p-1} r_i = p^{n-1}$.

Case 1. $r_0, r_1, \ldots, r_{p-1}$ all $\neq 0$.

Form the complete block C_1 (first member 0) by subtracting one element of A_1 from all other elements of A_1. If C_1 is multiplied by b we get the block C_2 formed also by subtracting one element of A_2 from all other elements of A_2. We can form the elements of C_2 (first member 0) by subtracting one of the r_0 elements of A_2 (first member 0) from itself and all the other $r_0 - 1$ such elements. Hence there are r_0 elements of C_2 (first member 0).

We can also form the elements of C_2 (first member 0) by subtracting one of the r_1 elements of A_2 (first member 1) from itself and all the other $r_1 - 1$ such elements. Hence there are r_1 elements of C_2 (first member 0).

Hence $r_0 = r_1 = r_2 = \ldots = r_{p-1} = p^{n-2}$.

This result must be true for all other first members, since all elements of the field are obtainable from those in A_1 by addition or subtraction.

Case 2. One or more of $r_0, r_1, r_2, \ldots, r_{p-1} = 0$.

Suppose in fact that $r_i, r_j, \ldots, r_k \neq 0$. Exactly as above, we can show $r_i = r_j = \ldots = r_k$. This leads to a contradiction since p^{n-1} is not divisible by a number less than p, unless we have

Case 3. All except one of $r_0, r_1, \ldots, r_{p-1} = 0$.

Suppose that the first members of all elements in A_2 are w, where $1 \leqslant w \leqslant p-1$. They cannot all be 0 since we can generate the whole field by addition and subtraction among the elements of A_1 and therefore the same among the elements of A_2. This would lead to all first members being 0 which is a contradiction unless b is the zero of the field.

Now we can find m such that $mw \equiv 1 \pmod{p}$. Hence mb (i.e. $b+b+\ldots+b$) times A_1 gives a block all of whose first members are also 1 (block D). Therefore multiplying by $(mb)^{-1}$ gives a block all of whose first members are 1 (block E). Subtract block D from block E (i.e. ith element from ith element) and obtain a block all of whose first members are 0. This must lead as before to first members of all blocks being 0. Hence the multiplier

$$(mb)-(mb)^{-1} = 00\ldots 0,$$

therefore $(mb) = (mb)^{-1}$, i.e. $mb = \pm$ the unity of the field.

Hence the only possible b for Case 3 are $00\ldots 0t$ where $1 \leqslant t \leqslant p-1$.

Write now the first members of the field elements in the order generated by c, a primitive root of the field, and its powers, i.e.

$$00\ldots00,\ 00\ldots01,\ c,\ c^2,\ \ldots,\ c^{p^n-2} \quad (C^{p^n-1}=1).$$

If these elements are multiplied in turn by $c, c^2, \ldots$, we obtain a cyclic permutation on all elements other than the zero, and by the above theorem any pair of columns satisfies the required symmetrical property. Multiplication by $c^{u(p^n-1)/(p-1)}$ will multiply columns by $00\ldots0t$, where t takes the values $1, 2, \ldots, p-1$, and hence the required property is satisfied only by the powers $c, c^2, \ldots, c^{(p^n-1)/(p-1)}$.

The proof that $c^{u(p^n-1)/(p-1)} = 00\ldots0t$ is as follows:

The element $00\ldots0t$ is expressible in the form C^x, therefore

$$c^{x(p-1)} = (00\ldots0t)^{p-1} \equiv 1 \equiv c^{p^n-1} \equiv c^{u(p^n-1)}.$$

Therefore $$x(p-1) = u(p^n-1).$$

For example, the elements of $GF(3^2)$ written in the order generated by powers of a primitive root are

00, 01, 11, 20, 21, 02, 22, 10, 12.

Taking the first members, we obtain a cyclic solution for $N=9$, $L=3$:

0	0	0	0
0	1	2	2
1	2	2	0
2	2	0	2
2	0	2	1
0	2	1	1
2	1	1	0
1	1	0	1
1	0	1	2

Thus, from the field $GF(p^n)$ we may obtain a cyclic solution for the case $L=p$, $N=p^n$. In the table of designs given below the first column of a cyclic solution is provided corresponding to the Galois fields 2^4, 2^5, 3^2, 3^3, 3^4, 5^2, 5^3 and 7^2. These have been taken from the tables in Stevens (1939), forming the basis of a series of completely orthogonal sets of cyclic Latin squares.

9. Experimental procedure

Suppose that the investigator is presented with an assembly containing 9 components and the problem of determining the effect of each of these in the performance of the whole. He decides upon an experiment in which each component appears at two values throughout and main effects are determined with a precision four times as great as that with which an assembly can be measured; in other words, the appropriate design is that for $L=2$, $N=16$. On referring to the table below he finds the design represented symbolically as follows (an explanation appears in a few lines):

$$++++-+-++--+---$$

The complete design is generated by taking this as the first column (or row), shifting it cyclically one place fourteen times and adding a final row of minus signs, thus:

```
+ − − − + − − + + − + − + + +
+ + − − − + − − + + − + − + +
+ + + − − − + − − + + − + − +
+ + + + − − − + − − + + − + −
− + + + + − − − + − − + + − +
+ − + + + + − − − + − − + + −
− + − + + + + − − − + − − + +
+ − + − + + + + − − − + − − +
+ + − + − + + + + − − − + − −
− + + − + − + + + + − − − + −
− − + + − + − + + + + − − − +
+ − − + + − + − + + + + − − −
− + − − + + − + − + + + + − −
− − + − − + + − + − + + + + −
− − − + − − + + − + − + + + +
− − − − − − − − − − − − − − −
```

The rows of this design may be taken as referring to assemblies and the columns to components. In the case in point there are nine components so that only nine columns are required. Select any nine columns, say the first nine, and obtain:

	Components								
	1	2	3	4	5	6	7	8	9
Assembly 1	+	−	−	−	+	−	−	+	+
2	+	+	−	−	−	+	−	−	+
3	+	+	+	−	−	−	+	−	−
4	+	+	+	+	−	−	−	+	−
5	−	+	+	+	+	−	−	−	+
6	+	−	+	+	+	+	−	−	−
7	−	+	−	+	+	+	+	−	−
8	+	−	+	−	+	+	+	+	−
9	+	+	−	+	−	+	+	+	+
10	−	+	+	−	+	−	+	+	+
11	−	−	+	+	−	+	−	+	+
12	+	−	−	+	+	−	+	−	+
13	−	+	−	−	+	+	−	+	−
14	−	−	+	−	−	+	+	−	+
15	−	−	−	+	−	−	+	+	−
16	−	−	−	−	−	−	−	−	−

The components have been labelled $1, 2, \ldots, 8, 9$: a plus corresponding to component 7 in assembly 3 means that in that assembly component 7 appears at its extreme value; a minus corresponding to component 3 in assembly 12 means that in that assembly component 3 appears at its nominal value; and similarly. It will be seen that each component appears eight times at an extreme value and eight times at nominal, so that the arrangement is perfectly symmetrical. The investigator now proceeds to set up assemblies according to this design, to measure whatever characteristic of them is in mind, and to record the results.

10. Analysis of the results

The results are in the form: measurement on assembly $1 = r_1$, measurement on assembly $2 = r_2$, ..., measurement on assembly $16 = r_{16}$. The effect of component 5, say, is required. Observe now that this component appears as plus in assemblies 1, 5, 6, 7, 8, 10, 12 and 13; and as minus in assemblies 2, 3, 4, 9, 11, 14, 15 and 16. Then the best estimate m_5 of the contribution of component 5 to the assembly characteristic due to its shift in value is

$$m_5 = (r_1+r_5+r_6+r_7+r_8+r_{10}+r_{12}+r_{13}-r_2-r_3-r_4-r_9-r_{11}-r_{14}-r_{15}-r_{16})/16,$$

all observations where component 5 appears as plus being taken positively and where it appears as minus being taken negatively, and the divisor being the number of assemblies

made up. A solution similar to this, and as simple, holds for all designs where $L = 2$, and the general method of which components to put in which assemblies and how to evaluate the effects should now be apparent.

The results provide in addition an estimate of the experimental error, obtained as follows. Suppose that instead of 9 components, 15 had been used, laid out in accordance with the experimental design given above. Then m_{12}, for example, would have been evaluated by the equation

$$m_{12} = (r_2+r_4+r_5+r_8+r_{12}+r_{13}+r_{14}+r_{15}-r_1-r_3-r_6-r_7-r_9-r_{10}-r_{11}-r_{16})/16.$$

In general, with n components, the quantities $m_{n+1}, m_{n+2}, \ldots, m_{4K-1}$ can be evaluated from the equations (number of assemblies $N = KL^2 = 4K$ here). Since there are just n components, these quantities should each be zero. In actual practice this will not be so due to experimental error. The variance due to error is estimated by the formula

$$s^2 = 4K(m_{n+1}^2+m_{n+2}^2+\ldots+m_{4K-1}^2)/(4K-n-1).$$

Here $s^2 = 16(m_{10}^2+m_{11}^2+\ldots+m_{15}^2)/6$ and the error variance of $m_i = s_i^2 = s^2/4K$. This formula is, as proved above, equivalent to the usual sum of squares of residuals divided by the degrees of freedom; degrees of freedom for error $= (4K-1)-n$.

A correction is necessary here. It will not usually be possible to select components whose values are exactly at nominal or extreme. All components will in any case have to be measured and the extent to which they differ from the aimed-at values will affect the values of m_i and s_i^2. Suppose that 'nominal' components are selected from a small range whose centre is the nominal value; and similarly at the extreme. For the ith component the difference in value between nominal and extreme is $2t_i$. If the component differs from the aimed-at value by c_i and if $b_i = c_i/t_i$, then the equations we are solving, instead of being of the form

$$r_j = M+a_{j1}m_1+a_{j2}m_2+\ldots+a_{jn}m_n,$$

where the coefficients a_{ij} are $+1$ or -1, are of the form

$$r_j = M+(a_{j1}+b_{j1})m_1+(a_{j2}+b_{j2})m_2+\ldots+(a_{jn}+b_{jn})m_n,$$

i.e. $R = (A+B)X$ where capital letters refer to the appropriate matrices. An approximate solution for X is obtained from $R = AX$, as above, and closer approximations may be obtained by iteration; a detailed treatment of the method is given in Lindley (1946).

Standard statistical methods now apply in determining the significance of effects and of differences between effects; whether the tolerance on a certain component may be increased and what would happen to the assembly characteristics if this were done; whether it is advisable to reduce the tolerance on another because of the large-scale effect allowed by the existing tolerance; whether the design of the assembly is correct in the sense that if both ends of the tolerance range have been explored the results show that the nominal value of each component is in the optimum position: these questions, and many like them depending on particular circumstances, may now all be answered. Errors may of course in all cases be reduced by replication, but it is suggested that, in order to obtain the best selection from the set of all possible assemblies (the complete factorial experiment) and thus minimize the errors due to interactions between components (here neglected as small), a complete design should be chosen in preference to the repetition of a smaller one. This aspect must not be confused with the fact that certain designs are obtained from smaller ones by the process of doubling, which is an entirely different thing. The designs in the table below (pp. 323, 324) will apply

directly to any experiment requiring less than 100 assemblies; should larger designs be required, they may be constructed by the general methods given.

11. RELATIONSHIP BETWEEN MULTIFACTORIAL AND BALANCED INCOMPLETE BLOCK DESIGNS

We begin for convenience with the definition of a balanced incomplete block design (Fisher & Yates, 1943). In this, v varieties are placed in blocks of k experimental units (k being less than v) such that every two varieties occur together in the same number (λ) of blocks; each variety appears r times in all and the number of blocks is b. Whence $rv = bk$ and

$$\lambda = r(k-1)/(v-1).$$

Consider any of the multifactorial designs for $L = 2$. Let the rows refer to blocks and the columns to varieties; and suppose that a plus sign represents the appearance of a variety in a block, a minus sign the non-appearance. For $N = 4m$ we obtain a balanced incomplete block design with $b = v = 4m-1$, $k = r = 2m$ and $\lambda = m$; the complementary design has $b = v = 4m-1$, $k = r = 2m-1$ and $\lambda = m-1$. The proof follows immediately from the orthogonality of the columns of the multifactorial design.

Now consider a complete multifactorial design F with N rows and L symbols; by complete we mean that the number of columns of F is $(N-1)/(L-1)$. Referring to § 4, suppose that all the elements of the diagonal matrix D are equal to $\sqrt{L}$, so that $Q'Q = L.I$. Let the rows of Q refer to the symbols $0, 1, 2, \ldots, L-1$. In the multifactorial design F using these L symbols replace each by the corresponding row of Q, omitting the 1 contributed by the first column. Add a first column of ones to the resulting matrix and obtain matrix A. Clearly $A'A = N.I$ and hence $AA' = N.I$. In any two rows of F a pair of unequal symbols in the same column contributes -1 to the scalar product of the corresponding rows of A; a pair of equal symbols contributes $+(L-1)$. Supposing that in these two rows of F there are λ pairs of equal symbols in the same column, and remembering the 1 at the beginning of each row of A, we have

$$1+(L-1)\lambda-1[(N-1)/(L-1)-\lambda] = 0,$$

whence

$$\lambda = (N-L)/L(L-1).$$

Let the rows of F refer to varieties and let each column represent L blocks, one corresponding to each of the L different symbols. By the result of the previous paragraph every two varieties occur together in the same number of blocks. We therefore obtain a balanced incomplete block design with parameters:

$$r = (N-1)/(L-1); \quad v = N; \quad b = rL; \quad k = N/L; \quad \lambda = (N-L)/L(L-1).$$

When $L = 2$, so that $N = 4m$, we can thus generate a large number of designs. When $L > 2$, we obtain balanced incomplete blocks with parameters

$$\begin{aligned} r &= 1+L+L^2+\ldots+L^{h-1}, \\ v &= L^h, \\ b &= L+L^2+L^3+\ldots+L^h, \\ k &= L^{h-1}, \\ \lambda &= 1+L+L^2+\ldots+L^{h-2}, \end{aligned}$$

where $L = p^m$, p a prime, and $h > 1$.

The balanced incomplete block designs formed from multifactorial designs, for which

$$r = (N-1)/(L-1); \quad v = N; \quad b = rL; \quad k = N/L; \quad \lambda = (N-L)/L(L-1);$$

are in fact of a special kind and have been called by Bose (1942) *affine resolvable*. A balanced incomplete block design is resolvable if we can separate the b blocks into r sets of n blocks each ($b = nr$) such that each variety occurs once among the blocks of a given set; and if in addition either (i) $b+1 = v+r$ or (ii) any two blocks belonging to different sets have the same number of varieties in common, then the other is true and the design is called affine resolvable because of its relation to certain finite Euclidean geometries. Bose has shown

(1) If a resolvable balanced incomplete block design is such that any two blocks belonging to different sets have the same number of varieties in common, then $b+1 = v+r$.

(2) If for a resolvable balanced incomplete block design $b+1 = v+r$ then any two blocks belonging to different sets have the same number of varieties in common. We have shown

(3) If a resolvable incomplete block design (i.e. one with r, v, b, k given but not necessarily balanced in the sense that every two varieties occur together in the same number of blocks) has $b+1 = v+r$ and is such that any two blocks belonging to different sets have the same number of varieties in common, then it is balanced.

To sum up, if a resolvable incomplete block design has any two of the following properties:

(i) any two blocks belonging to different sets have the same number of varieties in common,

(ii) balance,

(iii) $b+1 = v+r$,

then it has the third. The orthogonal matrix method we have used to prove (3) can also be used to provide short proofs of (1) and (2).

Consequently a multifactorial design can be formed from a balanced incomplete block design provided that the latter is resolvable with parameters

$$r = (N-1)/(L-1); \quad v = N; \quad b = rL; \quad k = N/L; \quad \lambda = (N-L)/L(L-1).$$

Bose has pointed out that affine resolvable designs can be constructed from the affine geometry $EG(h, p^m)$ (our notation) by taking varieties as points and blocks as $(h-1)$-flats; this construction gives all the multifactorial designs for $L > 2$ which have so far been obtained.

The most general aspect of the multifactorial design is obtained by considering each assembly as a block and each value of each component as a variety. We obtain a partially balanced incomplete block design (Bose & Nair, 1939) with parameters:

$$r = N/L; \quad v = L(N-1)/(L-1); \quad b = N; \quad k = (N-1)/(L-1);$$

$$\lambda_1 = N/L^2; \quad n_1 = [(N-1)/(L-1)-1]\,L; \quad \lambda_2 = 0; \quad n_2 = (L-1);$$

$$p^1_{ij} = \begin{bmatrix} [(N-1)/(L-1)-2]\,L & (L-1) \\ (L-1) & 0 \end{bmatrix}; \quad p^2_{ij} = \begin{bmatrix} [(N-1)/(L-1)-1]\,L & 0 \\ 0 & (L-2) \end{bmatrix}.$$

Although Bose & Nair state that general methods for the construction of partially balanced incomplete block designs are to appear, we have been unable to find them, so that this aspect of the multifactorial design does not yield more solutions of the problem.

12. Summary

Methods are developed to avoid the complete factorial experiment in industrial experimentation when the number of factors is so large that the standard procedure is impractic-

able. By assuming a simplified linear hypothesis, the problem of determining main effects with maximum precision is reduced to a combinatorial one. Practically all useful solutions of this have been found when each factor appears at two levels, but the solutions for more than two levels are fairly limited. The relationship of these solutions to some encountered in balanced incomplete blocks has been discussed.

We are indebted to Mr G. A. Barnard for suggesting the problem and the method of approach by least squares; and to Dr Bronowski for drawing our attention to the useful paper of R. E. A. C. Paley. The work was carried out as part of the programme of the Ministry of Supply, Research and Development (S.R. 17), and appears by kind permission of the Chief Scientific Officer.

TABLE OF DESIGNS

A. *Designs for* $L = 2$. The first row of any cyclic design is given opposite N, the number of assemblies. As stated at the end of § 4 (III) the matrix A here consists of plus and minus ones: these are denoted below by plus and minus signs. There is always a final row of minus signs —the basic assembly—to be added. In the designs for $N = 28, 52, 76, 100$ the square blocks are permuted cyclically amongst themselves; in the three latter cases the extra column has alternate signs throughout apart from the corner element. The larger designs are grouped in fives for convenience.

$N = 8$. + + + − + − −
$N = 12$. + + − + + + − − − + −
$N = 16$. + + + + − + − + + − − + − − −
$N = 20$. + + − − + + + + − + − + − − − − + + −
$N = 24$. + + + + + − + − + + − − + + − − + − + − − − −
$N = 28$. First nine rows

+ − + + + + − − −	− + − − − + − − +	+ + − + − + + − +
+ + − + + + − − −	− − + + − − + − −	− + + + + − + + −
− + + + + + − − −	+ − − − + − − + −	+ − + − + + − + +
− − − + − + + + +	− − + − + − − − +	+ − + + + − + − +
− − − + + − + + +	+ − − − − + + − −	+ + − − + + + + −
− − − − + + + + +	− + − + − − − + −	− + + + − + − + +
+ + + − − − + − +	− − + − − + − + −	+ − + + − + + + −
+ + + − − − + + −	+ − − + − − − − +	+ + − + + − − + +
+ + + − − − − + +	− + − − + − + − −	− + + − + + + − +

$N = 32$. − − − − + − + − + + + − + + − − − + + + + + − − + + − + − − +
$N = 36$. (Obtained by trial) − + − + + + − − − + + + + + − + + + − − + − − − − + − + − + + − − + −
$N = 40$. Double design for $N = 20$.
$N = 44$. + + − − + − + − − + + + − + + + + + − − − + − + + + − − − − − + − − − + + − + − + + −
$N = 48$. + + + + + − + + + + − − + − + − + + + − − + − − + + − + + − − − + − + − + + − − − − + − − − −
$N = 52$.

First eleven rows

+	+ − + − + − + − + −	+ − + − + − + − + −	+ − + − + − + − + −	+ − + − + − + − + −	+ − + − + − + − + −
+	− + − − − − − − − −	+ + − − − − + + + +	+ + + + − − + + − −	+ + − − + + − − + +	+ + + + + + − − − −
−	+ + − + − + − + − +	+ − − + − + + − + −	+ − + − − + + − − +	+ − − + + − − + + −	+ − + − + − − + − +
+	− − − + − − − − − −	+ + + + − − − − + +	− − + + + + − − + +	+ + + + − − + + − −	− − + + + + + + − −
−	− + + + − + − + − +	+ − + − − + − + + −	− + + − + − − + + −	+ − + − − + + − − +	− + + − + − + − − +
+	− − − − − + − − − −	+ + + + + + − − − −	+ + − − + + + + − −	− − + + + + − − + +	− − − − + + + + + +
−	− + − + + + − + − +	+ − + − + − − + − +	+ − − + + − + − − +	− + + − + − − + + −	− + − + + − + − + −
+	− − − − − − − + − −	− − + + + + + + − −	− − + + − − + + + +	+ + − − + + + + − −	+ + − − − − + + + +
−	− + − + − + + + − +	− + + − + − + − − +	− + + − − + + − + −	+ − − + + − + − − +	+ − − + − + + − + −
+	− − − − − − − − − +	− − − − + + + + + +	+ + − − + + − − + +	− − + + − − + + + +	+ + + + − − − − + −
−	− + − + − + − + + +	− + − + + − + − + −	+ − − + + − − + + −	− + + − − + + − + −	+ − + − − + − + + −

$N = 56$. Double design for $N = 28$.

$N = 60$. + + − + +   + − + − +   − − + − −   + + + − +   + + + − −   + + + + +   − − − − −   + + − − −   − + − − −
+ + − + +   − + − + −   − − + −

$N = 64$. Double design for $N = 32$.

$N = 68$. + + − − +   − + − − +   + − − − +   + + + − +   − + + + +   + + − − +   − − − + −   + + + − +   + − − − −
− − + − +   − − − − +   + + − − +   + − + − +   + −

$N = 72$. + + + + +   + + − + +   + − + − −   + + − + +   + − − − +   + − + − +   + − + − −   − + + + −   + − − + −
+ − − + +   + − − − +   − − + + −   + − − − +   − − − − −   −

$N = 76$.

+	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −
+	− +	− −	+ +	− −	− −	+ +	+ +	− −	+ +	− −	− −	− −	− −	+ +	+ +	+ +	− −	+ +	+ +	+ +	+ +	− −
−	+ +	− +	+ −	− +	− +	+ −	+ −	− +	+ −	− +	− +	− +	− +	+ −	+ −	+ −	− +	+ −	+ −	+ −	+ −	− +

+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −	+ −
+ +	+ +	+ +	− −	− −	− −	− −	+ +	− −	+ +	+ +	− −	− −	+ +	− −
+ −	+ −	+ −	− +	− +	− +	− +	+ −	− +	+ −	+ −	− +	− +	+ −	− +

The first three rows are given; to obtain the complete design the square blocks are permuted cyclically. The first column, apart from the corner element, has alternate signs.

$N = 80$. + + + − +   + − − + +   + + − + −   − + − + +   + + + + −   + + − − −   − + + − −   − + − + −   + − + − +
+ + − − +   + + + − −   + − − − −   − − + − +   + − + − −   − − + + −   − + − −

$N = 84$. + + − + +   − − + − +   + + + − −   − + + − −   − + − + −   + + + + +   + + − + −   − + + + −   + + − − +
− − − + +   − + − − −   − − − − +   − + − + +   + − − + +   + − − − −   + − + + −   − + −

$N = 88$. Double design for $N = 44$.

$N = 92$. This design has not yet been obtained.

$N = 96$. Double design for $N = 48$.

$N = 100$.

+ − + − + − + − + − + − + −	+ − + − + − + − + − + − + −	+ − + − + − + − + − + − + −	+ − + − + − + − + − + − + −
− + − − − − − − − − − − − −	+ + − − + + + + − − + + − −	+ + − − − − + + + + − − + +	+ + + + + + − − − − − − + +
+ + − + − + − + − + − + − +	+ − − + + − + − − + + − − +	+ − − + − + + − + − − + + −	+ − + − + − − + − + − + + −
− − − + − − − − − − − − − −	− − + + − − + + + + − − + +	+ + + + − − − − + + + + − −	+ + + + + + + + − − − − − −
− + + + − + − + − + − + − +	− + + − − + + − + − − + + −	+ − + − − + − + + − + − − +	+ − + − + − + − − + − + − +
− − − − − + − − − − − − − −	+ + − − + + − − + + + + − −	− − + + + + − − − − + + + +	− − + + + + + + + + − − − −
− + − + + + − + − + − + − +	+ − − + + − − + + − + − − +	− + + − + − − + − + + − + −	− + + − + − + − + − − + − +
− − − − − − − + − − − − − −	− − + + − − + + − − + + + +	+ + − − + + + + − − − − + +	− − − − + + + + + + + + − −
− + − + − + + + − + − + − +	− + + − − + + − − + + − + −	+ − − + + − + − − + − + + −	− + − + + − + − + − + − − +
− − − − − − − − − + − − − −	+ + − − + + − − + + − − + +	+ + + + − − + + + + − − − −	− − − − − − + + + + + + + +
− + − + − + − + + + − + − +	+ − − + + − − + + − − + + −	+ − + − − + + − + − − + − +	− + − + − + + − + − + − + −
− − − − − − − − − − − + − −	+ + + + − − + + − − + + − −	− − + + + + − − + + + + − −	+ + − − − − − − + + + + + +
− + − + − + − + − + + + − +	+ − + − − + + − − + + − − +	− + + − + − − + + − + − − +	+ − − + − + − + + − + − + −
− − − − − − − − − − − − − +	− − + + + + − − + + − − + +	− − − − + + + + − − + + + +	+ + + + − − − − − − + + + +
− + − + − + − + − + − + + +	− + + − + − − + + − − + + −	− + − + + − + − − + + − + −	+ − + − − + − + + − + − + −

+ − + − + − + − + − + − + −	+ − + − + − + − + − + − + −	+ − + − + − + − + − + − + −
+ + + + − − − − − − + + + +	+ + + + − − + + + + − − − −	+ + − − + + − − + + + + − −
+ − + − − + − + − + + − + −	+ − + − − + + − + − − + − +	+ − − + + − − + + − + − − +
+ + + + + + − − − − − − + +	− − + + + + − − + + + + − −	− − + + − − + + − − + + + +
+ − + − + − − + − + − + + −	− + + − + − − + + − + − − +	− + + − − + + − − + + − + −
+ + + + + + + + − − − − − −	− − − − + + + + − − + + + +	+ + − − + + − − + + − − + +
+ − + − + − + − − + − + − +	− + − + + − + − − + + − + −	+ − − + + − − + + − − + + −
− − + + + + + + + + − − − −	+ + − − − − + + + + − − + +	+ + + + − − + + − − + + − −
− + + − + − + − + − − + − +	+ − − + − + + − + − − + + −	+ − + − − + + − − + + − − +
− − − − + + + + + + + + − −	+ + + + − − − − + + + + − −	− − + + + + − − + + − − + +
− + − + + − + − + − + − − +	+ − + − − + − + + − + − − +	− + + − + − − + + − − + + −
− − − − − − + + + + + + + +	− − + + + + − − − − + + + +	+ + − − + + + + − − + + − −
− + − + − + + − + − + − + −	− + + − + − − + − + + − + −	+ − − + + − + − − + + − − +
+ + − − − − − − + + + + + +	+ + − − + + + + − − − − + +	− − + + − − + + + + − − + +
+ − − + − + − + + − + − + −	+ − − + + − + − − + − + + −	− + + − − + + − + − − + + −

B. *Designs for* $L = 3, 5, 7$. The first column is given below and the complete design is formed by permuting it cyclically $(N-1)/(L-1)-1$ times and adding a row of zeros. The corresponding orthogonal matrix A of § 4 (II) is obtained by replacing the component value symbols of the design by the rows of Q (§ 4 (I)) with its first column suppressed.

$N=9$, $L=3$. 01220211

$N=27$, $L=3$. 00101 21120 11100 20212 21022 2

$N=81$, $L=3$. 01111 20121 12120 20221 10201 10012 22021 00200 02222 10212 21210 10112 20102 20021 11012 00100

$N=25$, $L=5$. 04112 10322 42014 43402 3313

$N=125$, $L=5$. 02221 04114 13134 12021 10244 31402 00444 20322 32121 32404 22043 31230 40033 34014 41424 21430 34403
11241 03001 11302 33234 34231 01330 12243 2010

$N=49$, $L=7$. 01262 21605 32335 20413 11430 65155 61024 54425 03646 634

REFERENCES

BOSE, R. C. (1942). A note on the resolvability of balanced incomplete block designs. *Sankhyā*, **6**, 105.

BOSE, R. C. & KISHEN, K. (1940). On the problem of confounding in the general symmetrical factorial design. *Sankhyā*, **5**, 21.

BOSE, R. C. & NAIR, K. R. (1939). Partially balanced incomplete block designs. *Sankhyā*, **4**, 337.

FISHER, R. A. (1942). *The Design of Experiments*, 3rd ed. Oliver and Boyd.

FISHER, R. A. & YATES, F. (1943). *Statistical Tables for Biological, Agricultural and Medical Research*, 2nd ed., p. 14. Oliver and Boyd.

HOTELLING, H. (1944). Some improvements in weighing and other experimental techniques. *Ann. Math. Stat.* **15**, 297.

KISHEN, K. (1945). On the design of experiments for weighing and making other types of measurement. *Ann. Math. Stat.* **16**, 294.

LINDLEY, D. V. (1946). On the solution of some equations in least squares. *Biometrika*, **33**, 326.

PALEY, R. E. A. C. (1933). On orthogonal matrices. *J. Math. Phys.* **12**, 311.

STEVENS, W. L. (1939). The completely orthogonalized Latin square. *Ann. Eugen., Lond.*, **9**, 82.

Added in proof

(1) Since this paper was written, one of the authors (J. P. B.) has obtained a design for $L=3$, $N=18$, $n=7$, by trial and error. It is known that this is the largest value of n possible.

(2) The designs given above for $L=2$ provide what is effectively a complete solution of the experimental problem considered by Hotelling (1944) and Kishen (1945).

ON THE SOLUTION OF SOME EQUATIONS IN LEAST SQUARES

By D. V. LINDLEY

In carrying out an experiment of the above type it is often not possible to arrange that the factors occur at exactly their required values: they will deviate by a small amount in either direction from the ideal aimed at. It is possible to allow for this in the case of the analysis of the results of a two-level experiment.

The equations to be solved by least squares when the factors are at the ideal values are

$$y_i = M + \sum_j a_{ij} m_j \quad (i = 1, \ldots, N,\ j = 1, \ldots, n), \tag{1}$$

where the a_{ij}'s are plus or minus one and

$$m_j = \frac{\partial y}{\partial x_j} t_j,$$

where $2t_j$ is the difference between extreme and nominal, or what is usually half the tolerance allowed. This assumes that the origin is taken halfway between the extreme and the nominal so that the equations assume a more useful character. We can further suppose the extreme to be at a higher value of x_j than the nominal. Now suppose that there is a small deviation from the ideal e_{ij}, given by actual minus ideal, corresponding to each a_{ij}. If this deviation occurs at the extreme value, i.e. $a_{ij} = +1$, the new extreme value will be $t_j + e_{ij}$ from the origin: on the other hand, if it occurs at the nominal, $a_{ij} = -1$, the new nominal value will be $t_j - e_{ij}$ from the origin. So in either case the deviation is $t_j + a_{ij} e_{ij}$ from the origin.

So equations (1) should now read

$$\begin{aligned} y_i &= M + \sum_j a_{ij} \frac{\partial y}{\partial x_j}(t_j + a_{ij} e_{ij}) \\ &= M + \sum_j \left(a_{ij} + \frac{e_{ij}}{t_j}\right) \frac{\partial y}{\partial x_j} t_j \quad (\text{since } a_{ij}^2 = 1) \\ &= M + \sum_j \hat{a}_{ij} m_j, \end{aligned} \tag{2}$$

with $\quad \hat{a}_{ij} = a_{ij} + e_{ij}/t_j.$

Let us now write equations (1) in matrix notation

$$Y = AX, \tag{3}$$

i.e.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1n} \\ 1 & a_{21} & a_{22} & \cdots & a_{2n} \\ 1 & a_{31} & a_{32} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{N1} & a_{N2} & \cdots & a_{Nn} \end{pmatrix} \begin{pmatrix} M \\ m_1 \\ m_2 \\ \cdots \\ m_n \end{pmatrix}$$

Then equations (2) can then be written

$$Y = (A + B) X, \tag{4}$$

where B is the matrix
$$\begin{pmatrix} 0 & e_{11}/t_1 & e_{12}/t_2 & e_{13}/t_3 & \dots & e_{1n}/t_n \\ 0 & e_{21}/t_1 & e_{22}/t_2 & e_{23}/t_3 & \dots & e_{2n}/t_n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & e_{N1}/t_1 & e_{N2}/t_2 & e_{N3}/t_3 & \dots & e_{Nn}/t_n \end{pmatrix},$$
which is dependent on the deviations from the ideal values divided by *half* the difference between nominal and extreme.

Now the least square solution of (3) is easily found; it is
$$X_1 = (A'A)^{-1} A'Y.$$
We have, in order to solve (4), to solve the equations
$$(A+B)'\, Y = (A+B)'\,(A+B)\,X$$
or
$$A'Y + B'Y = A'AX + (A'B + B'A + B'B)\,X. \tag{5}$$

Since the elements of B are small in comparison with those of A, an approximate solution is provided by X_1, the solution of (3). If we put this in the smaller unknown term in (5) we get a second approximation in the solution of
$$A'Y + B'Y = A'AX + (A'B + B'A + B'B)\,X_1,$$
i.e.
$$X_2 = (A'A)^{-1}[A'Y + B'Y - (A'B + B'A + B'B)\,X_1],$$
in general the rth approximation is given in terms of the $(r-1)$th by
$$X_r = (A'A)^{-1}[A'Y + B'Y - (A'B + B'A + B'B)\,X_{r-1}].$$

This then enables successive approximations to the solution to be found, and we can carry it on until the accuracy is as great as we desire. This will usually be dictated by the accuracy with which the y_i and the e_{ij} were measured. In one practical case it was not found necessary to proceed beyond X_3. Since $(A'A)^{-1}$ is diagonal the solution at each stage is simple. Once $(A'B + B'A + B'B)$ has been calculated each stage only involves the computation of $(A'B + B'A + B'B)\,X_{r-1}$ and the subtraction of it from $A'Y + B'Y$. It is important to notice that B has a column of 0's and A a column of 1's corresponding to the mean M.

In the ideal experiment where the factors are all at nominal or extreme, the standard error s_k associated with m_k is given in terms of the residual s by
$$s_k^2 = (A_{kk}/D_A)\,s^2,$$
where D_A = the determinant of $A'A$, and A_{kk} = the minor of the (k,k)th element in $A'A$.

When $A'A$ is diagonal this is given by
$$s_k^2 = \frac{1}{c_{kk}}\,s^2,$$
where c_{kk} is the (k,k)th element of $A'A$.

In the practical case we then have
$$s_k^2 = (A+B)_{kk}/D_{A+B}\,s^2,$$
which to the first order in the e_{ij} is
$$s_k^2 = (A_{kk}/D_A)\,s^2 = \frac{1}{c_{kk}}\,s^2$$
as before.

Thus we have obtained without too large an amount of labour the solutions of the equations as accurately as we need and the standard errors of these solutions.

This work was carried out as part of the Research and Development programme of the Ministry of Supply (S.R. 17) and appears by permission of the Chief Scientific Officer.

SOME GENERALIZATIONS IN THE MULTIFACTORIAL DESIGN

By R. L. PLACKETT

1. The following remarks may be regarded as a sequel to a previous paper (Plackett & Burman, 1946), although the notation has been changed and compressed for convenience; it is hoped that no confusion thereby ensues. We consider first the determination of main effects, and find what modifications are required when certain of the orthogonal transformations previously used are no longer orthogonal.

2. We assume that y_r, the true value of the measurement on the rth assembly, is expressible in the form

$$y_r = \sum_A A_j \quad (r = 1, 2, \ldots, N;\ j = 1, 2, \ldots, a),$$

there being n components $A, B, \ldots, K$, where A_j is the effect due to component A at its jth value.

The vector $(A_1, A_2, \ldots, A_a)$ is denoted by $\mathbf{A}'$. Make the transformation $\mathbf{A} = A\mathbf{a}$, i.e.

$$A_i = \sum_j a_{ij} a_j,$$

to a new system of variables $(a_1, a_2, \ldots, a_a)$ represented by $\mathbf{a}'$; A is a non-singular $a \times a$ matrix, whose first column consists entirely of ones. Suppose A_i and B_j appear together in a design w_{ij} times; then A_i appears w_{i0} times where

$$w_{i0} = \sum_{j=1}^{b} w_{ij},$$

and similarly B_j appears w_{0j} times. For component B, $\mathbf{B} = B\mathbf{b}$, where $\mathbf{b}$ is a column vector and matrix B is a non-singular $b \times b$ matrix whose first column is all ones. With

$$M = a_1 + b_1 + \ldots + k_1,$$

we write

$$X' = (M, a_2, \ldots, a_a, b_2, \ldots, b_b, \ldots, k_2, \ldots, k_k),$$

and $Y = PX$, the first column of P consisting entirely of ones.

We find what conditions are satisfied if $P'P = N.I$. Within the columns corresponding to component A we have

$$\sum_i a_{ij} a_{ik} w_{i0} = N\delta_{jk}. \tag{1}$$

Again, considering the product of a column of component A with one of component B,

$$\sum_{i,k} a_{ij} w_{ik} b_{kl} = N\delta_{1j}\delta_{1l}. \tag{2}$$

In equation (1) put $j = 1$ and obtain

$$\sum_i a_{ik} w_{i0} = N\delta_{1k},$$

a set of linear equations which may be written

$$\sum_i a'_{ki} w_{i0} = N\delta_{1k},$$

the solution of which is

$$w_{i0} = \sum_k \bar{a}'_{ik} . N . \delta_{1k} = N\bar{a}_{1i}, \tag{3}$$

where $\bar{a}_{ij}$ is an element of matrix A^{-1}. Similarly $w_{0k} = N\bar{b}_{1k}$.

It follows immediately from equation (2) that

$$w_{ik} = N\bar{a}_{1i}\bar{b}_{1k}.$$

Therefore
$$Nw_{ik} = w_{i0}w_{0k}. \tag{4}$$

Finally, it is more convenient to express equation (1) in terms of the elements of A^{-1} rather than those of A. Suppose W is the matrix whose elements are $\delta_{ij}w_{i0}$. Then

$$A'WA = N.I, \quad \text{i.e.} \quad N.A^{-1}W^{-1}(A^{-1})' = I,$$

therefore
$$\sum_j \bar{a}_{ij}\bar{a}_{kj}Nw_{j0}^{-1} = \delta_{ik},$$

i.e.
$$\sum_j \bar{a}_{ij}\bar{a}_{kj}(\bar{a}_{1j})^{-1} = \delta_{ik}. \tag{5}$$

Condition (4) arises also in analysis of variance. If matrices A and B have orthogonal columns, then (3) becomes $w_{i0} = N/a$ and similarly $w_{0k} = N/b$, so that $w_{ik} = N/ab$. There is no difficulty in showing that conditions (3), (4) and (5) are sufficient for the validity of equations (1) and (2).

With designs and matrices satisfying these conditions we may therefore determine $a_2, a_3, \ldots, a_a, b_2, b_3, \ldots, b_b$ with maximum precision and our estimates of these parameters are independent. In particular cases, such non-orthogonal functions of the A_i and B_j may be of greater interest or moment than the orthogonal functions usually chosen. Two conclusions may thus be drawn:

(i) Having defined a set of linearly independent linear functions, not necessarily orthogonal, of the A_i, then these functions may be determined as independent statistics with maximum precision, provided (3), (4) and (5) are satisfied for all components present.

(ii) If, for any reason, a factorial or multifactorial design is incomplete owing to loss of observations, then provided (4) is satisfied it may be possible to find from (3) and (5) linear functions of the A_i which can be determined as independent statistics with maximum precision.

3. Designs for which $Nw_{ik} = w_{i0}w_{0k}$ may be constructed immediately from those given in Plackett & Burman (1946). Consider the design for N assemblies in which each component appears at L values. For such a lay-out $w_{i0} = w_{0k} = N/L$ and $w_{ik} = N/L^2$ $(i, k = 1, 2, \ldots, L)$. Corresponding to component A, divide the L symbols into groups of $u_1, u_2, \ldots, u_p$ so that

$$\sum_{i=1}^{p} u_i = L;$$

each member of a group is equal, so that the L symbols are successively replaced by $1, 1, \ldots, 1, 2, 2, \ldots, 2, \ldots, p, p, \ldots, p$. This transforms component A into one which appears at p values. Similarly, corresponding to component B we have groups $v_1, v_2, \ldots, v_q$ so that

$$\sum_{k=1}^{q} v_k = L.$$

We now have $w_{i0} = u_i.N/L$, $w_{0k} = v_k.N/L$, and $w_{ik} = u_i.v_k.N/L^2$. Thus $Nw_{ik} = w_{i0}w_{0k}$.

Condition (3) gives $u_i/L = \bar{a}_{1i}$ and $v_k/L = \bar{b}_{1k}$; a suitable value of L is now chosen so that u_i and v_k are all integers, and the design constructed.

4. We may extend the scope of our inquiry to include interactions, prefacing our extension by clarifying what appears to be a known result which gives orthogonal functions of

the observations corresponding to interaction degrees of freedom (Fisher, 1942). When first-order interactions are present the effect due to the ith value of component A and the jth value of component B is
$$R_{ij} = A_i + B_j + (AB)_{ij},$$
the term $(AB)_{ij}$ representing the interaction. We again make transformations $\mathbf{A} = A\mathbf{a}$ and $\mathbf{B} = B\mathbf{b}$ where the matrix A has a first column of ones and all columns are orthogonal to each other, similarly for matrix B. Consider first the quantities $(A_i + B_j)$. With $M = a_1 + b_1$ we can transform these into $(M, a_2, a_3, \ldots, a_a, b_2, b_3, \ldots, b_b)$, and the matrix R_1 of the transformation will consist of a first column of ones, followed by $(a-1)$ columns formed by repetitions of the rows of A, followed by $(b-1)$ columns which are repetitions of the rows of B. Clearly the columns of R_1 are mutually orthogonal. There remain $(a-1)(b-1)$ columns R_2 to be chosen so that the matrix $R = (R_1 \vdots R_2)$ is a square matrix with mutually orthogonal columns. Corresponding to the columns of R_2 we may choose quantities $r_{a+b}, r_{a+b+1}, \ldots, r_{ab}$ into which the $(AB)_{ij}$ may be transformed.

The columns of R_2 may be chosen arbitrarily, but there are two methods whereby they may be written down at once. The first method is the one referred to at the beginning of this section, which consists in taking the $(a-1)(b-1)$ inner products of a column of R_1 (not the first) belonging to component A with a column of R_1 (not the first) belonging to component B. Thus take the inner product of the tth and $(a-1+u)$th columns of R_1. The scalar product of this column with the vth column of R_1 $(2 \leqslant t,\ v,\ \leqslant a;\ 2 \leqslant u \leqslant b)$ is
$$\sum_{i,j} a_{it} b_{ju} a_{iv}.$$
Keeping i fixed and summing over j gives zero. Hence the columns of R_2 are orthogonal to those of R_1. That they are orthogonal between themselves follows similarly from the fact that
$$\sum_{i,j} a_{it} b_{ju} a_{iv} b_{jw} = 0 \quad \text{unless} \quad t = v \text{ and } u = w.$$

The second method is at present available only when $a = b = L = p^m$ (p a prime and m an integer). We refer to the modified factorial designs in § 7 of Plackett & Burman (1946). With matrix B equal to matrix A the symbols in the design for $N = L^2$ are replaced by the corresponding rows of matrix A, the first column of A being omitted.

Writing $ab = r$; $a_2, a_3, \ldots, a_a$, $b_2, b_3, \ldots, b_b$ respectively equal to $r_2, r_3, \ldots, r_{a+b-1}$; $\mathbf{R}$ the column vector elements R_{ij}; $\mathbf{r}$ the column vector elements $r_1, r_2, \ldots, r_r$; we now have $\mathbf{R} = R\mathbf{r}$ where R is a matrix, elements r_{ij}, whose first column consists of ones, all columns being mutually orthogonal.

5. Now regarding R as a single component and R_h $(h = 1, 2, \ldots, r)$ as the effect due to its hth value, we must have for maximum precision of determination of R_h and C_k, the effect due to component C at its kth value, that $w_{hk} = N/r.c$. Thus if w_{ijk} denotes the number of occurrences together of A_i, B_j and C_k, then $w_{ijk} = N/abc$. Hence for optimum determination of the effects A, B, (AB) and C, all values of components A, B, C, must appear together equally frequently. This condition, however, leads to the optimum determination of effects B, C, (BC) and A; also of effects C, A, (CA) and B; because any first-order interaction is connected with two components only, all combinations of values of which appear equally often with all values of the third. Hence if for any three components in a design, all combinations of values appear equally frequently, all first-order interactions as well as main effects may be determined with maximum precision. In order to estimate a particular interaction (AB), all combinations of values of A and B must appear equally often with the values

of any other component present; in which case the first-order interactions between A and any other component, and between B and any other component, are automatically determined. Generally, therefore, having decided on those interactions which are of interest, $N = K \times$ L.C.M. of all relevant triplets abc; when all components appear at L values each we obtain $N = KL^3$.

This result is immediately generalized and for interactions between $(t-1)$ components appearing with other components we must have $N = KL^t$ when each component appears at L values. As interactions of higher order are included, we obtain a whole series of designs building up to the complete factorial design or replications thereof.

6. Similarly, the results concerning non-orthogonal transformations may be extended. We use the same notation with respect to component R; the matrix R_1 is constructed in the same manner from matrices A and B, but no longer consists of orthogonal columns; and the columns of R_2 are no longer orthogonal. The same meaning is attached to w_{ijk} as in § 5 and w_{ij0}, w_{i00} are defined by analogy with § 2.

For orthogonality between the columns of the matrix formed by repetitions of the rows of R_1 and corresponding to the main effects of components A and B, we have $Nw_{ij0} = w_{i00}w_{0j0}$. If now the columns of R_2 are formed by inner products of pairs of columns of R_1 as in § 4, then a repetition of the argument used there, together with the condition that $Nw_{ij0} = w_{i00}w_{0j0}$, shows that in the complete design the columns of the matrix corresponding to interaction (AB) and formed by repetition of the rows of R_2 are orthogonal, both to columns corresponding to main effects A and B and between themselves.

From the orthogonality of the columns of component R to those of component C we deduce $Nw_{ijk} = w_{ij0}w_{00k}$ giving finally

$$N^2 w_{ijk} = w_{i00}w_{0j0}w_{00k}, \tag{6}$$

which again leads to the optimum determination of interactions (BC) and (CA). Further generalizations of condition (4) follow immediately, together with appropriate modifications of conclusions (i) and (ii) in § 2; conditions (3) and (5) must always be satisfied as they stand, apart from slight changes of notation necessary in condition (3).

7. We now give an illustration of the design of an experiment where main effects only of components are considered, but the transformations made on the $A_i, B_j, \ldots$ are not orthogonal. Suppose that each component appears at three values such that the difference between the high and medium values is twice the difference between medium and low (i.e. we might be considering resistors of value 2000, 4000 and 8000 ohms). The effects of low, medium and high values for component A are respectively A_1, A_2 and A_3, and we shall be interested in considering $2A_1 - 3A_2 + A_3$ which will be zero if the A_i are linear functions of resistance value. We shall also be interested in $A_3 - A_1$, measuring the total change in the measurement due to varying the resistance over the whole range. It should be remarked that even if component values are equally spaced (e.g. in our example the three resistor values were 2000, 4000 and 6000 ohms) we might wish to test the hypothesis that the effect of a shift from 4000 to 6000 ohms is twice that of a shift from 2000 to 4000 ohms: this would again involve us in testing whether $2A_1 - 3A_2 + A_3$ departed significantly from zero.

In order that all conditions may be satisfied we take:

$$\begin{aligned} a_1 &= xA_1 + yA_2 + zA_3, \\ a_2 &= -sA_1 + sA_3, \\ a_3 &= 2tA_1 - 3tA_2 + tA_3. \end{aligned} \quad \text{i.e. matrix } A^{-1} = \begin{bmatrix} x & y & z \\ -s & 0 & s \\ 2t & -3t & t \end{bmatrix}.$$

Applying condition (5) we obtain

$$x+y+z=1, \quad -2/x+1/z=0, \quad \text{i.e.} \quad x=2z.$$

Taking $x=\frac{1}{2}$, $y=\frac{1}{4}$, $z=\frac{1}{4}$ gives

$$A^{-1}=\begin{bmatrix} 1/2 & 1/4 & 1/4 \\ -1/\sqrt{6} & 0 & 1/\sqrt{6} \\ 2/\sqrt{48} & -3/\sqrt{48} & 1/\sqrt{48} \end{bmatrix}.$$

Thus $w_{10}=N/2$, $w_{20}=N/4$ and $w_{30}=N/4$, by condition (3). If we suppose that matrix B is identical with matrix A, we obtain for the matrix whose elements are w_{ij}, the number of coincidences of A_i and B_j, the following:

$$\begin{bmatrix} N/4 & N/8 & N/8 \\ N/8 & N/16 & N/16 \\ N/8 & N/16 & N/16 \end{bmatrix}.$$

With five components whose values are represented by 0, 1, 2 a design of the required type can be constructed according to the method of § 3 from the design for $N=16$, $L=4$ (as obtained by the methods of Plackett & Burman (1946)) by replacing the four symbols for the component values by 0, 0, 1, 2. This gives:

00000	00021	10102	20210
00000	00012	10201	20120
01111	01200	11020	21002
02222	02100	12010	22001

When designs of the complete factorial or multifactorial type are not so available the construction of such a design without requiring an excessive number of assemblies will often necessitate a certain amount of ingenuity.

It is perhaps of interest to remark on a certain transformation, sometimes made, which results in the matrix A consisting of a first column and a leading diagonal of ones, all other elements being zero. It will be clear from the foregoing analysis that in this case it is impossible in any design whatever to determine the a_i with maximum precision, and some alternative transformation should be used.

We conclude by pointing out that a large class of combinatorial problems has been raised, of which a comparatively small proportion may be solved by the methods so far evolved. Statements in the foregoing that an experimental design must take a certain form shou' not be taken as implying that the relevant combinatorial lay-out necessarily exists.

This paper is published as part of the programme of the Ministry of Supply, Research and Development (S.R. 17), appearing by permission of the Chief Scientific Officer.

REFERENCES

Fisher, R. A. (1942). *The Design of Experiments*, section 47, para. 3. Oliver and Boyd.

Plackett, R. L. & Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika* **33**, 305.

THE GROWTH, SURVIVAL, WANDERING AND VARIATION OF THE LONG-TAILED FIELD MOUSE, *APODEMUS SYLVATICUS*

By H. P. HACKER and H. S. PEARSON

II. SURVIVAL. By H. P. HACKER

CONTENTS

1. Introduction

The purpose of this paper is to bring together the data we have collected on the length of life under natural conditions of the long-tailed field mouse. The trapping, marking and releasing have been done only during winter months in order to avoid interference with breeding, and so to leave the population as intact as possible.

The work began in 1936–7 and was continued for six winters. The data on growth have been described by Hacker & Pearson (1944), and further papers on travel and variation projected. Some account of the disposition of traps and the amount of movement of the mice is necessary for a discussion of the evidence on the length of survival. A map of the area trapped is therefore included in this paper and will also apply to the more detailed account of the distances travelled which will be published later.

2. Method of marking mice

In the 1936–7 season we started by marking the right hind foot of each mouse with one of the metal rings used for identifying canaries. These were made of aluminium with numbers stamped on them; they fitted the limbs well but the metal was too soft. By rubbing on the ground and by being gnawed by the mice some of the numbers were made difficult to read, and the edges of some of the rings became sharp and jagged, injuring the mice so severely that we had to kill them with chloroform.

Evans (1942, p. 184) used rings, and when the foot swelled too badly for the ring to be removed he amputated the leg. He found that 'the majority of these individuals were subsequently recovered in later censuses and appeared to be in healthy condition'. There must be some doubt, however, whether such mutilated individuals should be included in survival records or analyses of travels.

After a trial of two weeks we gave up using rings and have not tried the nickel rings described by Chitty (1937, p. 41). He has kindly sent us samples of his rings for comparison and the metal is much harder than ours, but after our experience of puncturing ears as a method for marking mice we would not think of returning to ringing. One definite advantage of using metal rings is the possibility of recovering them from the voidings of predators and so tracing the fate of the mice. We once found a mummified mouse, and by soaking the ears in water could read its number, but it would be impossible to do this after the body had been eaten.

The marking of the ears is done without an anaesthetic with the animal held lightly in an assistant's hand, and it very rarely squeaks, bites or struggles after the first attempt to escape from the hand. If a mouse does bite or behave obstreperously, it is quite often found to have behaved in the same way before; for instance, one out of a family of six reared from birth maintained a reputation for biting whenever it was touched. The membranous ear of *Apodemus* seems therefore almost insensitive, and we have not had any reason to suspect that the punctures affect the life of the mouse in any way.

The instrument we use is a leather punch with an interchangeable die which makes holes of about 1·5 mm. in diameter. This is rather clumsy but costs only two shillings and is quite effective. A more elegant and efficient, but much more expensive, instrument is a dentist's rubber dam punch. The small spring punch which chicken rearers use is quite useless for this purpose. We tried several and did not get clean punctures; moreover, the slight click that this instrument makes is more disturbing to the mouse than the actual puncture.

The four quarters of the ear pinna give distinctive sites for puncturing, and by combinations of not more than three punctures in each ear more than 1000 different patterns can be made. By using each pattern twice, once for each sex, we could use the simpler patterns with only one or two punctures in each ear. These are easier to make and to identify, and we had enough to choose from without having to clip the toes as recommended by Burt (1940, p. 12). The lower anterior quarter of the pinna is more fleshy than the other three, so that punctures here tend to heal up and to need puncturing again when the mouse is recaught and examined. This difficulty can be largely overcome by making them as high up the margin of the ear as is possible without risk of confusion with an upper puncture.

Incidentally, we may remark that the method totally failed when applied to *Cletherionomys* (*Evotomys*), whose ears are short, hairy and thicker in texture. The punctures heal up readily, leaving puckered scars. We therefore did not try to mark individuals of this genus but merely made a single puncture on the left ear of all those caught, in order to show in any future trapping that it was not a new mouse. Even then a puckered left ear and a normal right ear was often the only sign that the mouse had possibly been caught before.

3. Method of trapping

We used the Selfridge trap described by Elton *et al.* (1931, p. 714), and by taking out two bars from the back added the nest box introduced for the Tring trap by Chitty (1937, p. 39).

The main disadvantage of this cheaper trap, the danger to the tail of the mouse, we tried to circumvent by fixing a small stop to prevent the door touching the floor when it shuts. But an injury to the tail is a minor disaster to the mouse, as it is well known that *Apodemus* can escape by shedding part of its tail (Barrett-Hamilton & Hinton, 1910–21, p. 502). Sumner & Collins (1918, p. 1) have some interesting records of this faculty among American species of mice. One of us found the skin of a tail at the mouth of a mouse hole, evidence of a narrow escape from a predator. What happens is that the skin slips off very readily and the exposed tendons and bone shrivel, or are gnawed off, leaving a stub. Some very remarkable deformities due to this and other injuries were found in mice caught for the first time, 57 out of 1000 such mice showing injuries not due to trapping, whereas 84 mice out of 1000 consecutive catches were found injured by the trap. As even very minor injuries were recorded among these with a view to their use in identification on a future occasion, the rate of injury, though regrettable, was not much greater than that which occurs normally in nature, and some of those we inflicted were observed to heal completely.

In passing we might record a difference in liability to accident noticed in getting out these figures. One male injured its tail three times out of the six times it was caught; on the other hand, a male of similar size was caught ten times without injury and a female fourteen times. A similar individuality to that noted in the matter of biting seems to be present, and it may be that the more 'cautious' ones tend to get their tails injured by being slower in entering the traps.

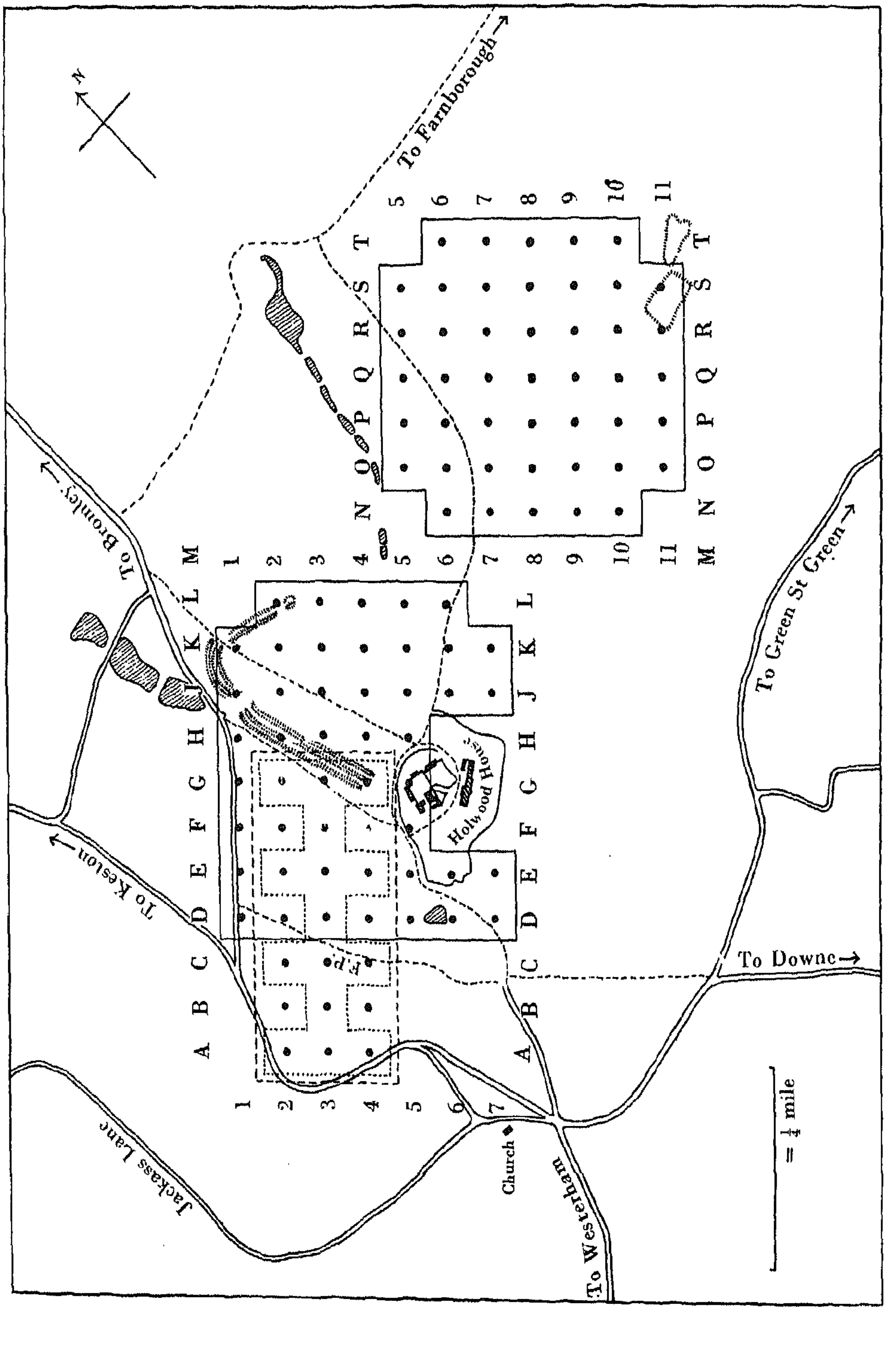
Undoubtedly the comfort and well-being of the captive mice is greatly increased and the number of deaths reduced by the nest box devised by Chitty. Their comfort is also increased (1) by putting the trap under cover of vegetation when possible, (2) by pointing the trap away from the prevailing wind and weather, and (3) by putting the trap on a slope so that water does not run into the nest box. The first two points probably add to the efficiency of the trapping as (1) the mice seem to like cover, and (2) sticks and leaves do not collect in the entrance and prevent the trap shutting.

A mechanical difficulty met with should be mentioned. With frequent rebaiting, the bar on which the bait hook hangs becomes so loose that the mouse can push it aside and escape. This can be prevented by the simple precaution of pushing the nest tin over the end of the bar to keep it fixed in position.

4. ARRANGEMENT OF TRAPS

In the 1936–7 season we worked at Studland in Dorset on the area that Diver (1933) has studied in such detail. Our main object was to find the distribution of mice in relation to the different types of habitat he had mapped out, and for this purpose we laid our traps at the most likely spots in each of the areas we were testing, without regard to the distance between these areas. We thus gained our own experience of how far *Apodemus* travels, and from our records there (which we hope to publish with maps) and from Chitty's (1937) useful summary of previous records of distances travelled by small mammals, we decided to use a grid of 100 yard squares in our routine trapping during the following winters. This we were able to do in Holwood Park, Keston, by permission of the late Lord Stanley and of Lady Stanley.

The reference map shows the part of the estate we used, with the grid of trapping centres dotted on it. The continuous line includes the two areas over which we trapped in 1937–8.



Map showing areas trapped in successive years and method of numbering trap sites. Distance between sites is 100 yd.

The western area consists of park and woodland, while the eastern area is mainly farmland, grazing and arable. The broken line is the rectangle, 700 by 300 yd., that we used in 1938–9, partly overlapping the wooded area trapped in 1937–8 and extending into the woods west of the public footpath (*F.P.* on map); as this area was too large to cover in one week the nine new sites in the west woods were trapped first and the twelve old sites in the east woods in the following week. The dotted line shows the area trapped in the next two years when the same middle strip of seven trapping centres was used with every alternate trapping centre on either side.

At each trapping centre we put out six traps in the form of a hexagon with sides of 10 yd., each trap site being also 10 yd. from the centre. Having marked the site we set the trap in the most favourable spot within a yard or so. Since neighbouring hexagons were trapped simultaneously it is not likely that a trapping centre would have caught mice living near another centre, but only those that lived in the intervening area. The hexagons on the edge of a trapped area have, however, a larger region from which to draw mice than hexagons in the interior, so that sometimes we shall describe results from the central and peripheral parts separately.

5. LENGTH OF TRAPPING PERIODS

In the 1936–7 season at Studland we gained the impression that if kept out for three or four nights a group of traps would catch most of the mice in the immediate neighbourhood, and that any mice caught later than this tended to come from a greater distance. An example of the kind of evidence on which this opinion was based may be quoted.

During nine nights in February eighteen traps were set in a hazel wood and the following are the numbers of mice caught each night:

9, 3, 1, 0, 0, 4, 1, 4, 0.

They show a rapid decrease from nine on the first day to none at all on the fourth and fifth nights, and then a renewal of catches suggesting outside mice coming into the area. The same impression is gained from Table 1. The relation between the day on which a mouse was caught and the distance it lived from the traps is further dealt with in § 13.

In the 1937–8 season we used 3 days as the routine period. In 1938–9 we increased this to 4 days in most months but to 5 days in November in the east woods, and to 6 days in March and April in both woods. In later years routine trapping was only done in December and March, and we increased the time to as many days as seemed convenient or necessary, our longest time being 10 days.

Although the catch of the first day or two may not always be good, as in March 1941 (Table 1), the results usually confirmed our choice of 4 days for the minimum period of trapping, as in March 1940 (Table 2).

The chief reason for this difference in the rate of catching was undoubtedly the weather. We can be fairly sure that a moonless or cloudy night is favourable to wandering, a wild night of south-west wind and rain seeming as good as any.* These conditions are perhaps unfavourable to owls. Snow and possibly also hoar frost seem to be unfavourable conditions. We did not keep weather records but only notes of striking changes in the weather. If later on we can compare our catches with the nearest meteorological record we may be

* Burt (1940, p. 25) came to much the same conclusion about the effect of weather, and Evans (1942, p. 190) emphasizes the effect of abnormally wet weather in increasing the number caught.

able to speak with more certainty about the effect of weather. But it will be seen in the following section on the efficiency of trapping that most of the mice that we have reason to expect to be in the neighbourhood do get caught, and that the length of trapping we adopted is justified.*

One point should be borne in mind when considering these records. The mice were kept in cages and only set free, when the trapping was finished, at the centre of the trapping site at which they were caught. This leaves the area vacant and free for outside mice to come in. Rather different results would be obtained if the mice were set free at once and allowed to stay in their home area, and it is certain that under these conditions a good many more traps would be needed as a reserve to remove the local mice each successive night. Chitty warned us that local mice block the traps each night and our own experiences confirm this. For example, we set free a series of thirty-one females as soon after they were caught as possible because they were either pregnant or nursing mothers. These mice gave an aggregate of ninety-nine captures, all on consecutive nights except for one mouse that missed two nights and one that missed one night.

Table 1. 24–29 *March* 1941. *Six traps in each group. Six groups for* 6 *days*

Index no. to site (see map)	Catch on successive nights
B2	1, 5, 5, 3, 1, 2
D2	1, 5, 4, 3, 0, 4
F2	0, 0, 6, 6, 0, 0
B4	2, 4, 7, 3, 1, 3
D4	3, 4, 2, 3, 0, 1
F4	1, 5, 5, 3, 0, 1
Totals	8, 23, 29, 21, 2, 11

Table 2. 18–23 *March* 1940. *Six traps in each group. Four groups for first* 4 *days, and four groups for* 6 *days*

Index no. to site (see map)	Catch on successive nights
A2	5, 5, 0, 0
C2	5, 2, 0, 0
A4	4, 2, 3, 0
C4	2, 3, 1, 0
D3	3, 2, 1, 0, 0, 0
E3	3, 2, 0, 0, 0, 0
F3	6, 1, 0, 0, 0, 0
G3	3, 7, 0, 0, 0, 0
Totals	31, 24, 5, 0, 0, 0

The tables show the catch from all traps set during the periods stated. Each group of six traps is entered separately to show the local variation in rate of catching that occurred. The index number for each group enables a reader to find its position on the map.

6. Efficiency of the trapping

In order to know how definitely we can assume that a mouse had either died or emigrated because it was not caught at any given trapping, it is necessary to get some idea of the likelihood of catching any mouse known to be alive in the area, in other words to have some test of the efficiency of the trapping. The best figures for this test are shown in Table 3; they are those for the season 1938–9, in which seven trappings were made in each hexagon throughout the area marked by the broken line on the map. Each horizontal line in the table represents a batch of mice all caught for the first time and for the last time in the

* Bole (1939, p. 57) has given reasons for choosing 3 days as the period for trapping.

same two months, which are indicated by the ×'s at either end of the line.* At each intervening trapping the whole batch must have been alive but might have missed being caught; the table shows for each batch:

(*a*) the number of possible catches (bold type);

(*b*) the number of mice missed (italics).

The numbers (*a*) and (*b*) are totalled for all batches at the bottom of the tables, and hence the percentage of misses can be calculated for each month.

Table 3. *Monthly distribution of misses*, 1938–9

East woods

No. of mice	Months N.	D.	J.	F.	M.	A.	J.
15	×	**15**	**15**	**15**	**15**	**15**	×
	×	*10*	*0*	*3*	*0*	*0*	×
17	×	**17**	**17**	**17**	**17**	×	
	×	*7*	*0*	*6*	*0*	×	
5	×	**5**	**5**	**5**	×		
	×	*1*	*0*	*1*	×		
5	×	**5**	**5**	×			
	×	*2*	*0*	×			
9	×	**9**	×				
	×	*5*	×				
5		×	**5**	**5**	**5**	**5**	×
		×	*0*	*2*	*0*	*0*	×
2		×	**2**	**2**	×		
		×	*0*	*0*	×		
1		×	**1**	×			
		×	*0*	×			
12			×	**12**	**12**	**12**	×
			×	*6*	*3*	*1*	×
8			×	**8**	**8**	×	
			×	*4*	*2*	×	
9			×	**9**	×		
			×	*4*	×		
2				×	**2**	**2**	×
				×	*0*	*0*	×
1				×	**1**	×	
				×	*0*	×	
4					×	**4**	×
					×	*0*	×

Summary

Month	D.	J.	F.	M.	A.	Grand total
Total possible catches	**51**	**50**	**73**	**60**	**38**	**272**
Total misses	*25*	*0*	*26*	*5*	*1*	*57*
Percentage of misses	49	0	36	8·3	2·6	21·0

West woods

No. of mice	Months N.	D.	J.	F.	M.	A.	M.
12	×	**12**	**12**	**12**	**12**	**12**	×
	×	*0*	*0*	*1*	*0*	*1*	×
4	×	**4**	**4**	**4**	**4**	×	
	×	*1*	*1*	*0*	*0*	×	
2	×	**2**	**2**	**2**	×		
	×	*0*	*0*	*0*	×		
2	×	**2**	×				
	×	*0*	×				
15		×	**15**	**15**	**15**	**15**	×
		×	*1*	*0*	*0*	*0*	×
2		×	**2**	**2**	**2**	×	
		×	*0*	*0*	*0*	×	
4		×	**4**	×			
		×	*0*	×			
13			×	**13**	**13**	**13**	×
			×	*2*	*2*	*0*	×
1			×	**1**	**1**	×	
			×	*0*	*0*	×	
3			×	**3**	×		
			×	*0*	×		
2				×	**2**	**2**	×
				×	*1*	*0*	×
2					×	**2**	×
					×	*0*	×

Summary

Month	D.	J.	F.	M.	A.	Grand total
Total possible catches	**20**	**39**	**52**	**49**	**44**	**204**
Total misses	*1*	*2*	*3*	*3*	*1*	*10*
Percentage of misses	5·0	5·1	5·8	6·1	2·3	4·9

Note on East Woods. For December and February 51 misses out of 124 or 41·1 %; for other months 6 misses out of 148 or 4·1 %.

* Thus of the fifty-one mice caught in the east woods for the first time in November, nine were caught for the last time in January, five for the last time in February, five in March, seventeen in April and fifteen in June.

In the east woods there were many misses in December and February owing to snow in the week of trapping, but if these two records are excluded the remaining figures (including no miss in January) give a total proportion of misses of only 4·1 %. In the west woods, where no exceptionally bad weather occurred during trapping, the total proportion of misses was 4·9 %.

If the figures of Table 3 are subdivided between (1) central trapping sites surrounded by other trapping sites, and (2) peripheral trapping sites on the edge of the trapped area, we find the following results:

East Woods

Central	21	misses	out of	93	chances	or	22·6 %
Peripheral	36	,,	,,	179	,,	,,	20·1 %
Total as in table	57	,,	,,	272	,,	,,	21·0 %

West Woods

Central	3	misses	out of	64	chances	or	4·7 %
Peripheral	7	,,	,,	140	,,	,,	5·0 %
Total as in table	10	,,	,,	204	,,	,,	4·9 %

There is obviously no real difference in the incidence of misses in the centre and periphery of the area. The only distinction we have been able to detect is that the six mice that missed twice running were all on the periphery, and even this does not mean much as there were only five trapping sites in the centre compared with sixteen on the periphery.

It thus appears that under ordinary conditions of weather the rate of misses was about 5 %, or that there was a 20 to 1 chance of catching a given mouse. In bad weather the chance of catching was less, but missing a mouse in two consecutive months was rare. We may conclude therefore that the trapping was effective and that a mouse no longer caught had probably either died or emigrated.

Another opportunity for testing the efficiency of trapping occurred in the next season, 1939–40. Only two trappings, in December and March, were made over the whole area, but four hexagons were also trapped in February (D 3, E 3, F 3 and G 3). This smaller trapping was done in a short mild spell in the phenomenally severe frost of that year, to find out how many of the December mice had survived the hard weather. Out of twenty mice caught in both December and March not one was missed in February. Perhaps hunger helped to cause this complete catch.

7. Monthly survival rate

In considering the records of survival for 1938–9 (details of which are given in § 8), all mice accidentally killed have been excluded for obvious reasons. In the present section all mice when caught for the first time have also been excluded because so many were found to disappear within the first month after capture; these are studied separately in § 12.

The proportions surviving each month, shown in Table 4, are so similar as to suggest that the rate of disappearance during the 4 months considered was very nearly constant and can be approximately described by the mean monthly survival rate 0·876. This ratio, which means that we should expect seven out of eight mice to be alive at the end of a month, can be used as a standard with which to compare survival over other periods of time and in other groups of mice. For this purpose we may use the formula $y_x = y_0\,(0{\cdot}876)^x$, where x is the time in

months measured from the beginning of the period, y_0 the number of mice at the beginning of the period and y_x the number at the end.

Since the number of mice disappearing is proportional to the number present, either the predators disappear at the same rate, or they find increasing difficulty in catching the mice as these become larger and scarcer; possibly the larger the mouse the more lasting the meal.

Table 4. *Number of mice surviving from one trapping till the next. December* 1938 *to April* 1939

Period	Mice caught in first month	No. surviving at next trapping	Proportion surviving
Dec.–Jan.	82	72	0·878
Jan.–Feb.	105	91	0·867
Feb.–Mar.	141	124	0·879
Mar.–Apr.*	131	108	0·824* (*0·879*)

* This period was 1·5 months and the proportion for one month is shown in italics. This rate r is calculated thus: $108/131 = r^{1\cdot 5}$.

8. SURVIVAL FROM ONE TRAPPING TO THE NEXT

The seven trappings of 1938–9 enable us to record the progressive diminution of the group of mice first caught in any given month. It will be seen from the map described on p. 337 that only fifteen of the hexagons set in this season were set again in the following year, so that the mice from the other six hexagons are not recorded as their later history is not known. Trapping stretched over a fortnight, as we could not cover the whole area in a week; the western seven hexagons were set in the first week and the eastern eight in the second.

Trapping started on 14 November 1938, and the four intervals between the first five trappings were each of 1 month. The interval between the March and April trappings was 1·5 months (6 weeks). Then the western part was trapped in May after an interval of 1 month, and the eastern part in June after 2 months. These two trappings are combined to form the seventh trapping at an average interval of 1·5 months, giving an error of 2 weeks which is negligible over the whole period. The area was trapped again in December, giving an interval of about 7 months, and the next interval was exactly 3 months to March 1940. Not one of these mice was caught after this although fifteen hexagons were trapped in the following December and March.

These details as to times, and those already given about the places and methods of trapping, are tedious, but are essential for estimating the reliance that can be placed upon the results obtained. The facts relating to the times of trapping are used for calculating the 'periods of survival' used in the following tables, and up to the sixth trapping may be used for finding the actual date of trapping if necessary.

Eighty-nine mice were caught in November, and the number of these caught again or known to be alive at each successive trapping is shown in Table 5. The results from the middle strip of seven hexagons are shown separately from those for the eight marginal hexagons; the two are then added together to give the totals for the eighty-nine mice.

The numbers first caught, forty-five and forty-four, in the two subdivisions of the area, are so similar that the figures are readily comparable without calculating proportions, and it is obvious that there is no appreciable difference between the margin and the centre of the area. Below the table the numbers expected from the standard survival rate are compared with the actual totals. The number for the first month is much lower than the standard, and this will be discussed in § 12 on the meaning of mice caught once only. During the rest of the winter the figures correspond closely, but after the April trapping the totals fall below the numbers expected. The significance of the difference between observed and expected survivals is indicated under each figure by the ratio of this deviation to the standard

Table 5. *Survival of mice first caught in November* 1938

Month of trapping	Nov. 1938	Dec. 1938	Jan. 1939	Feb. 1939	Mar. 1939	Apr. 1939	May–June 1939	Dec. 1939	Mar. 1940
Length of survival in 'lunar months'	0	1	2	3	4	5·5	c. 7	c. 14	c. 17
Mice from middle row of hexagons	45	29	27	22	20	18	13	2	1
Mice from marginal hexagons	44	27	24	21	19	16	7	2	1
Totals	89	56	51	43	39	34	20	4	2
Nos. expected with monthly survival rate of 0·876	*89*	*78*	*49*	*45*	*38*	*32*	*28*	*8*	*3*
Deviation / *Standard error*	*0*	*−7·05*	*0·79*	*−0·71*	*0·62*	*0·85*	*−3·51*	*−1·79*	*−0·07*

error according to the formula $\frac{n_t - pn_{t-1}}{\sqrt{[n_{t-1}p(1-p)]}}$, where $p = 0{\cdot}876$, n_{t-1} is the number of mice known to be alive during the preceding month, and n_t the number found to be alive the following month.

Another series worth studying are the seventy-three mice caught in January 1939. Table 6 shows how this group is made up of twenty-five mice remaining out of thirty first caught in December, and forty-eight first caught in January.

Those caught first in January show the low survival for the first month noticed in Table 5. The twenty-five mice first caught in December were survivors of a group of thirty, a small catch owing to snow in the east woods. In this case there was no excessive loss in the first month, and the fact that such a loss did not occur when only a few mice, presumably living near the trap sites, were caught has a bearing on the discussion of mice caught once only (see § 12).

After the May–June trapping the totals are again below the numbers expected, so that the survival rate of the winter generation appears to have diminished as the summer generation took its place. This lowering of survival rate for the summer months is of doubtful significance in Tables 5 and 6, since for reasons discussed elsewhere, the May–June numbers are not very reliable (p. 341). A more reliable figure for survival over 8·5 months,

including the summer, can be obtained if the disappearance by December 1939 of all the mice caught in April is considered. The April catch consisted of eighty-six mice, including the seventy-three shown as surviving in April in Tables 5 and 6. In December only eight of these eighty-six were recaught, whereas for the 8·5 months under consideration the 0·876 rate would give an expected number of 27·9 mice. The ratio of deviation from expectation to the standard error is 10·7, which is highly significant.

Table 6. *Survival of mice first caught in December* 1938 *and in January* 1939

Month of trapping	Jan. 1939	Feb. 1939	Mar. 1939	Apr. 1939	May–June 1939	Dec. 1939	Mar. 1940
Length of survival in lunar months	0	1	2	3·5	*c.* 5	*c.* 12	*c.* 15
Mice first caught in December	25	21	17	16	14	1	0
Mice first caught in January	48	37	32	23	16	0	0
Totals	73	58	49	39	30	1	0
Numbers expected with the standard monthly survival rate of 0·876	*73*	*64*	*51*	*40*	*32*	*12*	*1*
Deviation / Standard error	*0*	*−2·11*	*−0·72*	*−0·43*	*−0·82*	*−4·05*	*−1·43*

9. SURVIVAL FROM ONE YEAR TO THE NEXT

We have also a series of long-term observations and can record the proportion of mice surviving from one season to another. Here again the number of survivors can be compared with the number that would have survived if the standard monthly survival rate had been effective.

(*a*) From 1937–8 to 1938–9. Of the areas trapped during the first of these seasons (see p. 337 and Map) the farm land was not re-trapped in the second season, nor was the greater part of the park and woods, so that no long-term observations are available from the mice caught in these areas except that they were not found as migrants elsewhere.

In the rectangle of 400 by 300 yd. trapped in both seasons thirty-eight mice were caught during the first season. Of these, twenty-one were proved to be alive in March when the first season ended, but not one of them was caught during the second season, either in November when trapping began or in any of the six subsequent trappings in each of the hexagons throughout the area. With the standard survival rate, six out of the twenty-one should have been alive in November.

(*b*) From 1938–9 to 1939–40. It will be seen from § 8 (p. 341) that only fifteen of the hexagons trapped in the first of these two seasons are available for this comparison. Here 218 mice were caught during the first season, and of these 137 were known to be still alive in March. Of this group only eight were caught in the December trapping of the second season and only three in March. Not one was found in the 1940–1 season, although the whole area was then trapped twice. The numbers to be expected from the standard survival rate are thirty-six for December and twenty-four for March.

(c) From 1939–40 to 1940–1. In both these seasons identical areas were trapped in December and March (see p. 337 and Map). The figures are:

Total caught in 1939–40 season		134	
No. known to be still alive in March 1940		79	
No. of these known to be alive in the following season:			
	In December 1940	2	(Expected 21)
	In March 1941	1	(Expected 14)

The survival from year to year is thus seen to be much less than that which the standard monthly survival rate based on the winter months of 1938–9 would lead us to expect (see § 7). Survival rates calculated for the three winter months December to March of 1939–40 and 1940–1 are 0·757 and 0·815 respectively (see data in Table 7). These rates are lower than the standard rate of 0·876, but, as will be seen later (p. 347), this is to be expected, since mice caught once only are included. The corresponding summer rate for March to December 1940 is 0·693, as only two mice survived out of seventy-nine. As in 1939 this is again lower than the winter rate.

10. Survival in relation to size and sex

Table 7 *a*, *b* and *c* shows the proportion of mice surviving over 3 months interval in three successive seasons. The weights at the first trapping are divided into three main groups: (1) below 12·5 g., (2) from 12·5 to 19·9 g. and (3) 20 g. and over. Males are shown in bold type and females in italics. Proportions have been calculated only for the totals of each weight group, as the numbers of mice in the small divisions are so few. Among the winter mice shown in this table the weights of the males and females cover the same range, so that they can be combined in the same table for the advantage of larger numbers, although the weight groups are not really equivalent for the two sexes. The separate rates for each sex can be seen from the table, and perhaps the main error introduced by combining the sexes is that the females of 17·5 to 19·9 g. should be included in the over 20·0 g. group if that group is to be regarded as the fully grown mice.

None of the proportions in the table is convincing by itself, but the uniformity seen throughout the three seasons seems to show that the survival rate is lower for very small and very large mice than for those of intermediate weight.

The following analysis, to which the χ^2 test has been applied, shows that there is no significant difference in the survival rate of the sexes:

From Table 7 *a*

Males	37 survivors	out of 68,	proportion	=0·54
Females	26 „	„ 48,	„	=0·54
Totals	63 „	„ 116,	„	=0·54

From Table 7 *b*

Males	23 survivors	out of 49,	proportion	=0·47
Females	19 „	„ 48,	„	=0·40
Totals	42 „	„ 97,	„	=0·43

From Table 7 *c*

Males	43 survivors	out of 75,	proportion	=0·57
Females	36 „	„ 71,	„	=0·51
Totals	79 „	„ 146,	„	=0·54

Table 7. *The survival of mice over three winter periods of 3 months each: (a) From November 1938 to February 1939; (b) From December 1939 to March 1940; (c) From December 1940 to March 1941. Males are shown in bold type and females in italics*

	Weights when first caught							
	Below 12·5 g.		From 12·5 to 19·9 g.			20 g. and over		
	7·5–	10·0–	12·5–	15·0–	17·5–	20·0–	22·5–	25·0–
	(*a*) From November 1938 to February 1939							
Totals caught:								
Males and females	**10** *7*	**12** *5*	**13** *12*	**10** *9*	**4** *2*	**11** *6*	**6** *4*	**2** *3*
Survivors	**2** *3*	**6** *3*	**10** *9*	**7** *4*	**2** *1*	**6** *3*	**3** *2*	**1** *1*
Both sexes:								
Totals caught	17	17	25	19	6	17	10	5
Survivors	5	9	19	11	3	9	5	2
Totals for each weight group:								
Survivors	14 out of 34		33 out of 50			16 out of 32		
Proportion surviving	0·41 ± 0·08		0·66 ± 0·07			0·50 ± 0·09		
	(*b*) From December 1939 to March 1940							
Totals caught:								
Males and females	**1** *0*	**1** *2*	**8** *12*	**16** *18*	**18** *8*	**1** *6*	**2** *2*	**2** *0*
Survivors	**0**	**0** *1*	**3** *5*	**8** *8*	**10** *4*	**0** *1*	**1** *0*	**1**
Both sexes:								
Totals caught	1	3	20	34	26	7	4	2
Survivors	0	1	3	16	14	1	1	1
Totals for each weight group:								
Survivors	1 out of 4		38 out of 80			3 out of 13		
Proportion surviving	0·25 ± 0·22		0·48 ± 0·06			0·23 ± 0·12		
	(*c*) From December 1940 to March 1941							
Totals caught:								
Males and females	. .	**3** *7*	**24** *30*	**26** *23*	**15** *7*	**6** *2*	**1** *2*	. .
Survivors	. .	**0** *4*	**14** *18*	**16** *9*	**10** *3*	**3** *1*	**0** *1*	. .
Both sexes:								
Totals caught	.	10	54	49	22	8	3	.
Survivors	.	4	32	25	13	4	1	.
Totals for each weight group:								
Survivors	4 out of 10		70 out of 125			5 out of 11		
Proportion surviving	0·40 ± 0·15		0·56 ± 0·04			0·45 ± 0·15		

Standard errors of proportions are calculated by formula $\sqrt{\frac{pq}{n}}$.

As has been remarked on p. 344, the proportions surviving are lower than 0·67, which is the proportion expected if the standard monthly survival rate of 0·876 had been maintained throughout the 3 months.

11. Discussion of data on survival

The progressive diminution of the population described in §§ 7–10 may be due to: (1) mice learning to avoid traps, (2) emigration from the area, (3) death.

(1) If the mice learned to avoid traps in any degree, we would expect chance catches with long periods of absence from the traps. We have seen, in § 6 on the efficiency of trapping, how rare such records are. There were only six mice missed from the traps on two or more successive trappings, and these were found on the edge of the trapped area where occasional visitors from the outside might be expected.

(2) We found no evidence of migration of mice during the winter months in which the trapping was done. Detailed evidence on this subject will be given when the records of travels are described in a later paper.

(3) Thus, although the first two causes cannot be excluded, it is likely that the main cause of the disappearance of the mice is their death. In captivity mice can survive for longer periods than those recorded at Holwood; one of a family of newly born mice that we kept in a cage lived for four years and five months.* That the Holwood mice appeared to survive for only a short part of their possible life was probably chiefly due to predators. There are many enemies in the Holwood area: the dejecta of owls have been found containing remains of mice, there are badger and fox earths, stoats and grass snakes have been caught, while the cats and dogs from neighbouring houses must also take their toll.

It has been seen that the survival rate of the winter population is much lower during the summer. This may be due to the tendency, already detected in winter, for the largest mice to die out (p. 344), since by April the surviving population is almost entirely composed of large mice (Hacker & Pearson, 1944, p. 159). Very small winter mice were also found to have a low survival rate, and should this hold good for the new season's young any very great increase in the population would be checked. The year to year fluctuations in the size of the population are probably closely connected with weather conditions affecting the length of the breeding season. Hacker & Pearson (1944, p. 161) have shown the effect of the early spring and late autumn of 1938 on the constitution of the population. The most favourable condition for its increase would appear to be a late autumn followed by an early spring, in which the survivors of a large winter population start breeding early. If such a condition recurred a 'plague' year might result (Elton, 1942), but of this we have had no experience in Holwood. On the other hand, a long winter might be expected to lead to a dearth of mice.

12. The large disappearance of mice in the first month after capture

In § 8 we have studied the gradual disappearance of two large series of mice caught in November 1938 and January 1939. Certain points of interest arise from studying the survival of each month's catch of new mice; a larger number of mice is available, as all the twenty-one hexagons trapped in 1938–9 can be used instead of the fifteen in § 8.

* This mouse still showed signs of an epiphyseal line at the lower end of the femur, a condition well known in Muridae.

Table 8 shows the results obtained in the two parts of the woods lying to the east and west of the footpath. The interval between each trapping was 4 weeks as described in § 8. Read horizontally the table shows the number surviving at each successive trapping out of the batch of mice caught each month. The ratio of survival is shown in brackets after each figure and this can be compared with 0·876, the standard monthly survival rate.

The proportion surviving for 1 month of all mice caught for the first time is 172 out of 238, or 0·723, and for all other catches is 287 out of 328, or 0·875, approximately the standard rate. The rate is consistently lowest for each new batch throughout the table except for February when only three mice were caught; and this is why new mice have been omitted in calculating the standard survival rate. A discussion of the reasons for this large number of mice caught once only is necessary before considering the origin of the new mice which continue to be caught each month.

Table 8. *Monthly survival of the batch of mice caught each month*

Month of first trapping	Nos. of mice caught in the first month and surviving in the following months				
	Nov.	Dec.	Jan.	Feb.	Mar.
East woods:					
Nov.	83	58* (0·70)	52 (0·90)	44* (0·85)	38 (0·86)
Dec.	—	13	9 (0·69)	8 (0·89)	7 (0·88)
Jan.		—	45	34 (0·76)	28 (0·82)
Feb.			—	3	3 (1·00)
Mar.				—	14
West woods:					
Nov.	32	24 (0·75)	20 (0·83)	18 (0·90)	18 (1·00)
Dec.	—	35	24 (0·69)	21 (0·88)	17 (0·81)
Jan.		—	21	16 (0·76)	16 (1·00)
Feb.			—	6	4 (0·67)
Mar.				—	6

* The actual figures were 55 and 43; these have been adjusted to make allowance for probable misses in December and February.

(*a*) The more rapid disappearance of very large and very small mice described in § 10 must have some effect in the earlier part of the trapping season when we have shown them to be more common (Hacker & Pearson, 1944, p. 161).

(*b*) The small catch in December and February in the east woods was due to snow and the evidence in § 6 on the efficiency of trapping showed that fifty-one out of the fifty-seven failures to catch in this area occurred in these months. This would lower the survival rate for November and January respectively, as some of the mice would have been recorded for the last time in those months instead of in December and February and so increase the number of mice caught once only. The November catch in the west woods was also below expectation, thirty-two mice compared with thirty-five in December; this may have been due to the traps being used for the first time after a treatment with linseed oil, and would have a slight effect in the same direction.

(*c*) Mice living at some distance from the traps, and caught after the local mice have been removed, may have died owing to difficulty in getting back to their homes when all

the mice are set free simultaneously at the centre of their hexagon or trapping site. We have seen on p. 337 some evidence that such a draining-in of outsiders does occur, and it is probable that these outsiders are at a disadvantage when set free together with the local mice. Data on the problem of how far mice travel from their homes, and from what distance they find their way back will be given in a later paper. To detect the actual home of any mouse is almost impossible; when mice were set free we often watched their behaviour and saw them disappear into holes, but these are as likely to be temporary shelters as their homes. On the other hand mice that are consistently caught at one trapping site in the first day or two of the trapping probably live near that site, therefore we shall now study the day on which mice were caught in each trapping period.

Table 9. *East woods. Day of catching in each month of three groups of mice*

Stay-at-homes caught every month							Stay-at-homes that missed being caught at least once							Travellers						
Sex	Months					Mean day	Sex	Months					Mean day	Sex	Months					Mean day
	N.	D.	J.	F.	M.			N.	D.	J.	F.	M.			N.	D.	J.	F.	M.	
M.	1	1	1	1	1	1·0	F.	1	.	1	1	1	1·8	M.	3	1	1	1	2	1·6
M.	1	1	1	1	1	1·0	F.	1	.	1	2	2	2·2	M.	2	2	1	1	2	1·6
F.	1	1	1	1	2	1·2	M.	3	.	1	1	2	2·4	M.	3	.	3	1	1	2·6
M.	1	1	1	1	2	1·2	F.	2	1	1	.	3	2·4	M.	4	.	2	3	1	3·0
M.	1	1	1	1	2	1·2	F.	4	.	1	1	2	2·6	M.	4	.	1	4	1	3·0
M.	3	1	1	1	2	1·6	F.	5	1	1	.	3	3·0	M.	4	.	2	1	3	3·0
M.	3	1	1	1	2	1·6	M.	3	3	1	.	3	3·0	M.	2	.	3	3	3	3·2
M.	2	2	1	2	1	1·6	F.	2	2	2	.	5	3·2	M.	4	.	1	.	3	3·6
F.	1	4	1	1	1	1·6	F.	3	.	3	4	2	3·4	M.	4	.	1	.	5	4·0
M.	3	1	1	2	2	1·8	F.	2	.	3	4	5	3·8	M.	5	.	2	.	3	4·0
F.	4	1	1	2	2	2·0	M.	4	.	2	.	3	3·8	M.	5	.	1	.	5	4·2
F.	2	2	2	2	2	2·0	M.	2	.	3	.	5	3·8							
M.	4	4	1	2	2	2·6														
Mean	2·0	1·6	1·1	1·4	1·6	1·5	Mean	2·6	4·0	1·6	3·4	2·8	2·9	Mean	3·6	4·4	1·6	3·1	2·6	3·1

13. Day caught and distance from traps

It can be seen from Table 8 that fifty-six mice lasted throughout the period from November to March. These may be divided into *Travellers*, defined as mice caught in more than one trapping site, and *Stay-at-homes*, the mice caught at one site only. The latter can be further divided into (1) those caught in each month of the period, and (2) those which missed being caught at least once. In comparing travellers with stay-at-homes only east woods' mice are taken, since eleven out of the twelve travellers were caught in the east woods and since trapping was not simultaneous in the two areas so that the day of catching may have been influenced by different weather.

Table 9 shows the day on which each east woods' mouse was caught in each of the 5 months. A dot indicates that a mouse missed being caught, in which case the mouse is regarded as having been caught on the fifth day in calculating the mean day of catching; the traps were only left out on a fifth day in November and March in the east woods and in March in the west woods.

It will be seen that among the mice caught in every month the mean day for each mouse ranges from 1·0 to 2·6, while for all the mice it is 1·5; none was caught as late as the fifth day. Among the travellers the mean day ranges from 1·6 to 4·2, and the grand mean is 3·1. The stay-at-homes that missed being caught at least once have a range and grand mean resembling those of the travellers, among which there was also a high proportion of misses.

In the west woods there was no snow in December and February, and only three mice missed being caught in all the five months, too small a sample to be worth recording. Table 10 shows the stay-at-homes caught in each month for comparison with those from the east woods which they closely resemble in the distribution of day of catching.

Table 10. *West woods. Day of catching of stay-at-homes caught every month*

Sex	Months					Mean day
	Nov.	Dec.	Jan.	Feb.	Mar.	
M.	1	1	1	1	1	1·0
F.	1	1	1	2	2	1·4
M.	3	1	1	1	1	1·4
F.	3	1	1	1	1	1·4
F.	3	1	1	1	1	1·4
M.	4	1	1	1	1	1·6
M.	4	1	1	2	1	1·8
M.	1	1	1	3	4	2·0
F.	2	1	2	3	2	2·0
M.	1	3	2	3	2	2·2
F.	4	2	1	2	2	2·2
F.	4	2	2	1	3	2·4
F.	4	2	3	3	4	3·2
M.	3	4	3	4	4	3·6
Mean	2·7	1·6	1·5	2·0	2·1	2·0

The means for the stay-at-homes caught each month are correct values, but the true means for the other groups might have been higher had the trapping been indefinitely continued, since some at least of the missed mice are likely to have been caught later than the fifth day allotted to them; the real difference between the groups may therefore have been greater than shown.

Instead of comparing means, the frequency of each day of catching in the four groups of mice can be shown by histograms as in Fig. 1; these illustrate the differences just described.

The travellers were clearly mice that lived within wandering distance of more than one trapping site and therefore probably at a greater distance from any one of these than the stay-at-homes, so that they tended to be caught on a late day after the stay-at-homes had been removed. The large number of misses suggests that in the unfavourable weather of December and February they did not wander far enough to be caught at all. The stay-at-homes which also missed in those months also tended to be caught on a late day and probably also lived at some distance from the traps. These data support the assumption that the day of catching is an indication of the distance at which a mouse lives from the traps.

From Tables 9 and 10 it appears that the day of catching was considerably influenced by the month, a reflexion of seasonal and fortuitous weather conditions (p. 337). The monthly means for the different sets of mice are given at the foot of the tables and are compared in Fig. 2. November and March were on the whole late months and January an early month; the lateness of the travellers in December and February is due to the large number of misses (counted as fifth day catches) in these months.

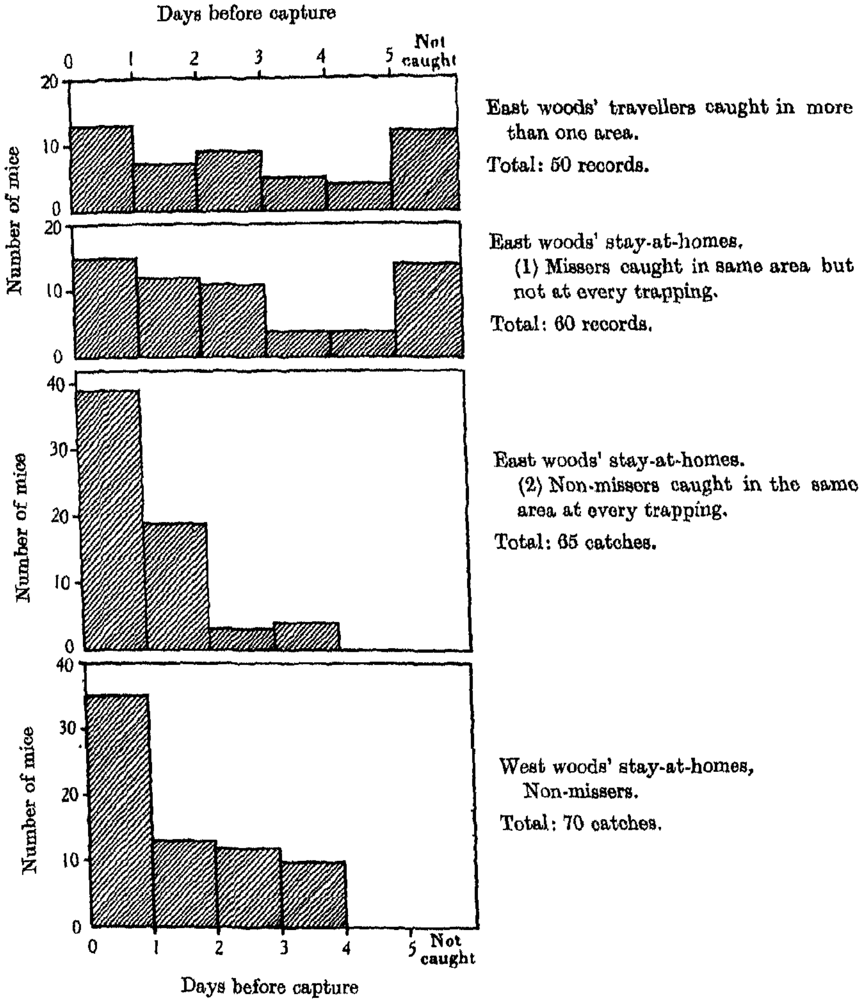


Fig. 1. Frequency of day of catching (from Tables 9 and 10).

Owing to this source of variability in the day of catching it is clear that in considering the mice caught once only these must not all be grouped together to obtain a mean day, since many more were caught in some months than in others. A frequency table of these mice is given in Table 11, and if the distribution of the days of catching and the monthly means at its foot are compared with those of the stay-at-homes from the corresponding parts of

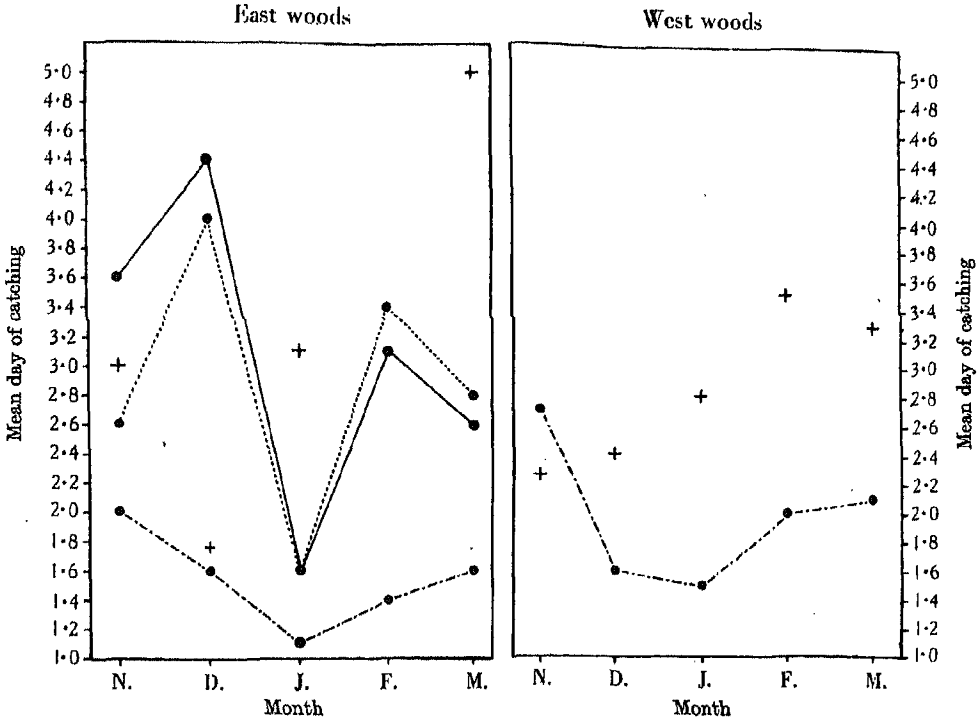


Fig. 2. Data from Tables 9, 10 and 11. Mean day of catching in each month for:
Stay-at-homes caught every month — · — · — · —
Stay-at-homes missed at least once ....
Travellers, visiting more than one site ————————
Mice caught once only +

Table 11. *Mice caught once only. Number caught on each day of trapping in the two parts of the woods*

East woods						West woods					
Day of catching	No. of mice					Day of catching	No. of mice				
	Nov.	Dec.	Jan.	Feb.	Mar.		Nov.	Dec.	Jan.	Feb.	Mar.
1	4	3	1	0	0	1	3	3	0	0	0
2	8	0	3	0	0	2	2	3	1	0	1
3	6	0	1	0	0	3	1	3	4	1	0
4	3	1	6	0	0	4	2	2	0	1	2
5	7	.	.	.	5	5	.	.	.	.	0
Mean day	3·0	1·75	3·1	.	5·0	Mean day	2·25	2·4	2·8	3·5	3·3

the woods, as is done in Fig. 2, it will be seen that on the whole they tend to be later. This tendency for the mice caught once only to resemble the travellers in being caught on a late day indicates that they too lived for the most part at a distance from the traps.

14. Day caught and frequency of catching

The relationship of day caught to frequency of catching may next be considered from the point of view of a single trapping site, irrespective of whether a mouse was caught in other months in any other trapping site or not. The travellers of the last section are regarded as having missed being caught, but misses are not counted as fifth day catches and any actual fifth day catches are omitted, which makes it permissible to combine east and west woods and so obtain larger numbers. To make use of further available data, the April catch is included throughout this section.

Twenty mice were always caught at the same trapping site in the first 4 days of trapping in each of the 6 months November 1938 to April 1939. The following figures show that they tended to be caught on the first day:

Day of catching	1	2	3	4	Total
Number of catches	65	32	11	12	120
Percentages	54	27	9	10	100

Mean day of catching 1·75.

In Table 12 these percentages are combined with those for mice caught at one trapping site in only 5, 4, 3, 2, 1 of these months, and caught elsewhere, or on the fifth or sixth day of trapping, or not at all, in the other months. The number of mice in each group is given in col. 2. This number multiplied by the number in col. 1 will give the total catchings on which the percentages are based. Thus the actual numbers can be reconstituted and the mean day of catching, shown in col. 4, calculated.

Table 12. *Percentage of catches on each day of trapping*

(1) No. of months mouse caught in one locality	(2) No. of mice	(3) Day of catching				(4) Mean day of catching
		First	Second	Third	Fourth	
6	20	54	27	9	10	1·75
5	31	53	26	13	8	1·75
4	34	39	27	21	13	2·07
3	41	32	27	22	19	2·29
2	39	21	24	28	27	2·62
1	140	22	24	30	24	2·56

The percentage of first day catches is seen to decrease markedly, and that of third and fourth day catches to increase, as the number of times a mouse was caught in the same locality decreases. The mean day of catching for each group of mice shown in the last column also increases, but the significance of this figure is considerably reduced by the monthly variation due to weather conditions already noted in the last section; for if mice are grouped together simply on the grounds of the number of times they were caught

throughout the whole season, it is clear that the different months will not be evenly represented in each group. In Table 13 the monthly means are therefore shown separately, revealing the differences.

In spite of these monthly differences, except in November, the tendency of the mean day to become later among the mice less frequently caught is still evident if each month's array is considered separately. Fig. 3 shows this diagrammatically.

The data are evidence that the more often mice are caught in any locality the earlier on the whole was the day of catching, which gives further support to the assumption that this day indicates how far a mouse lived from the traps.

Table 13. *The mean* day of catching for each month in mice grouped according to the number of times caught*

No. of times mouse caught in one locality	Months						Grand mean
	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	
6	2·5	1·8	1·3	1·8	1·9	1·4	1·75
5	2·2	1·8	1·4	1·6	2·2	1·6	1·75
4	2·3	1·7	2·0	2·2	2·3	1·8	2·07
3	2·3	1·9	2·1	2·6	2·7	1·7	2·29
2	2·8	2·2	2·3	3·2	2·9	2·6	2·62
1	2·45	2·3	3·0	3·3	3·0	2·2	2·56
Mean	2·4	1·9	2·1	2·3	2·4	1·9	

* The number of mice on which these means are calculated can be obtained from Fig. 3.

15. SURVIVAL RATE AND FREQUENCY OF CATCHING

Since the change in day of catching is a gradual one, and there is little difference in the tables between mice caught once only in any one locality and mice caught twice only, the question arises whether the survival rate also varies with the number of times a mouse was caught. If this were so, the mice which were only proved alive over a short period would seem as likely to have failed to revisit the traps through living at a distance from them as to have failed to survive. If a return is made to Table 8 and the mice are grouped according to how long they were known alive in the area as a whole, the survival rate of each group can be calculated and compared with the standard rate of 0·876.

No. of months previously known alive	No. of mice	Survivors next month	Survival rate
0	238	172	0·717
1	165	145	0·879
2	101	86	0·851
3	62	56	0·903

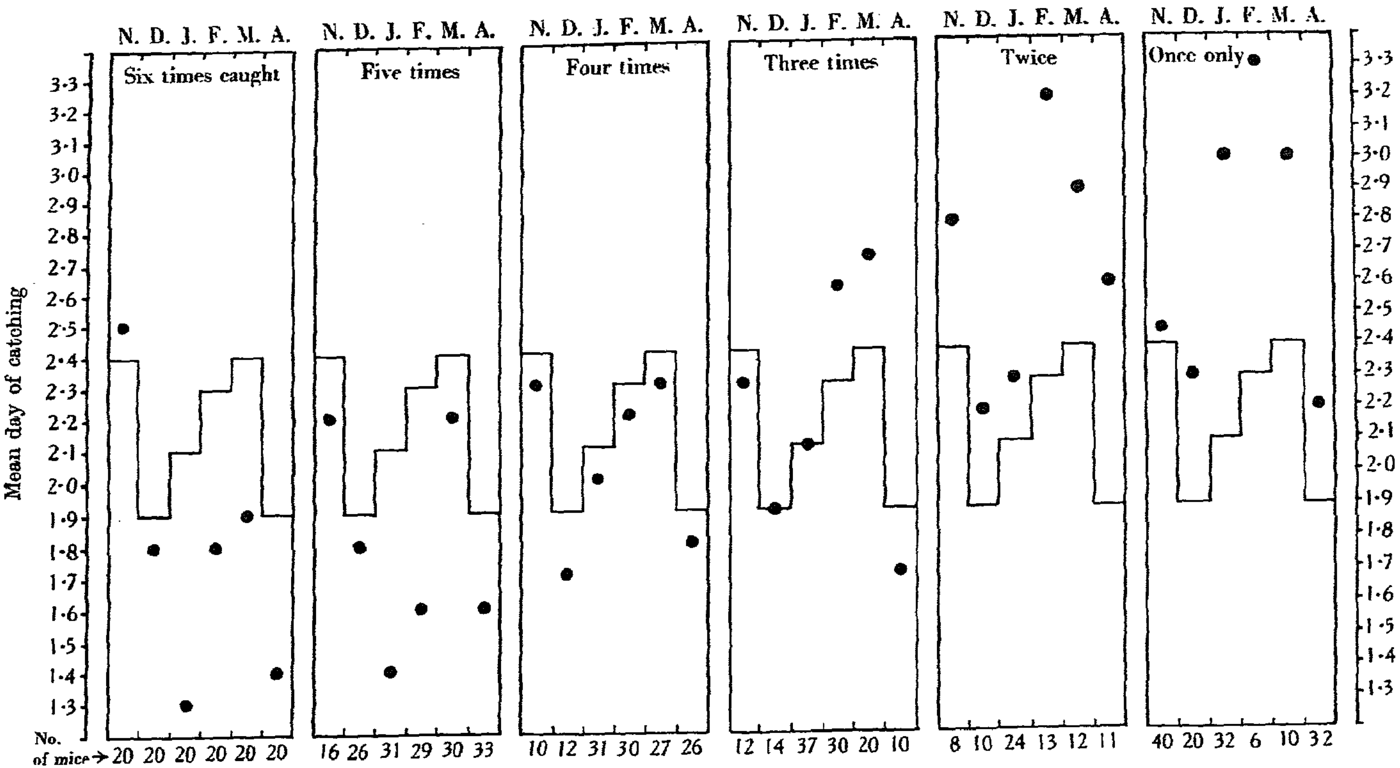


Fig. 3. Relation of mean day of catching to the monthly mean day in each of the six groups of mice in Table 13. The mean day of catching for all mice in each month, given at the bottom of Table 13, is shown by the horizontal step; these mean lines are the same in each of the rectangles. *The dots represent* the data from the body of that table and show that the mean day of catching becomes progressively later as the number of times the mice were caught decreases, until for mice caught only once or twice all the dots are above the line.

Here there is no gradual increase in the survival rate comparable with the gradual decrease in the mean day of catching; but the survival rate of mice caught for the first time is outstandingly low. From this it may be inferred that whereas mice were being drained into the traps over a continuous area there was a limit to the distance from which they could find their way home; those living beyond this limit failed to reach home under the conditions of the experiment and probably died. This would account for the lowering of the survival rate of first catches, as suggested on p. 347, and justifies the exclusion of these mice in calculating the standard rate in § 7.

16. Relation between disappearance of mice and appearance of new mice

In § 12, Table 8 was read horizontally to trace the survival of each batch of new mice. If the columns are added up vertically we get the total number, newcomers and old acquaintances, caught each month. This has been done and the results are shown in col. 5

Table 14

Month caught	No. caught		Summary		No. caught each month (5)
	New mice (1)	Mice lost (2)	Total caught to date (3)	Total lost to date (4)	
East woods:					
Nov.	83	—	83	—	83
Dec.	13	25	96	25	71
Jan.	45	10	141	35	106
Feb.	3	20	144	55	89
Mar.	14	12	158	68	90
Apr.	7	21	165	89	76
West woods:					
Nov.	32	—	32	—	32
Dec.	35	8	67	8	59
Jan.	21	15	88	23	65
Feb.	6	10	94	33	61
Mar.	6	6	100	39	61
Apr.	7	10	107	49	58

No. caught each month, east and west woods combined:
Nov. 115, Dec. 130, Jan. 171, Feb. 150, Mar. 151, Apr. 134.

of Table 14. Cols. 1 and 2 are readily extracted from Table 8, and the figures are summed up to date in cols. 3 and 4 from which the monthly figures in col. 5 may also be derived by taking the total number lost to date from the total caught to date. The low catches for November in the west woods and December in the east woods have been referred to on p. 347. The latter was made up for by the large catch in January, the average of the two months being nearly the same as that for the whole period.

The striking feature of this table is the apparent stability of the population as judged by the number caught each month (see col. 5 and foot of the table); although in each month and at each trapping site the new mice did not exactly replace those lost, in the area as a whole and over the whole period they more than replaced them.

Light on this problem can be obtained by studying the proportion of new mice to old caught on each successive day of the trapping. The data are given in detail in Table 15. Although the material is the same as that studied in Tables 8 and 14 the figures do not agree for two reasons. In studying survival a mouse is considered to be alive if it is caught in a subsequent month, but in this table only actual captures are entered. On the other hand mice excluded from the tables of survival because they were accidentally killed in the trapping are included here, as the point of interest is whether they had been caught before or not.

Table 15

Month caught	First day				Second day				Third day				Later				Total caught			
	New		Old		New		Old		New		Old		New		Old		New		Old	
	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.
East:																				
Dec.	5	5	11	10	0	1	2	3	1	0	4	1	2	0	2	1	8	6	19	15
Jan.	4	2	20	16	5	7	12	6	3	6	5	3	13	5	0	2	25	20	37	27
Feb.	0	0	17	11	0	1	3	6	0	0	8	2	2	0	6	6	2	1	34	25
Mar.	0	0	11	3	1	1	12	13	2	0	14	5	9	1	10	4	12	2	47	25
Apr.	2	0	14	15	1	0	11	7	2	1	6	2	0	1	9	3	5	2	40	27
West:																				
Dec.	6	3	8	6	7	7	4	4	8	3	1	0	3	4	2	0	24	17	15	10
Jan.	2	0	12	14	2	3	8	6	6	3	5	2	3	2	0	0	13	8	25	22
Feb.	0	0	11	11	0	0	6	6	2	1	11	5	3	2	8	1	5	3	36	23
Mar.	0	0	12	3	1	0	4	8	0	0	3	3	4	1	13	8	5	1	32	22
Apr.	0	0	18	9	1	1	6	7	1	0	6	2	2	2	4	2	4	3	34	20
Totals:																				
New	19	10			18	21			25	14			41	18			103	63		
Old			134	98			68	66			63	25			54	27			319	216

From an inspection of the columns it is obvious that there are more males than females except on the second day. The proportion of males in the new mice is 103 out of 166, or 62 %. This may not represent the condition in the field but indicate that the males are drawn from a greater area than the females. It is known from the work of Chitty (1937, p. 52), Burt (1940, p. 25), Blair (1942, p. 27) and others, that males of *Apodemus* and *Peromyscus* travel more widely than females. Some evidence of this can also be found in Tables 9 and 10, although the groups are small. All the eleven travellers there are males but only half of the stay-at-homes, fourteen out of twenty-seven in the east woods and seven out of fourteen in the west woods. If however the fourteen east woods' stay-at-homes which missed being caught in bad weather are considered as a separate group, ten are found to be females, while nine of the eleven male travellers also missed in bad weather. This suggests that males travelled further than females but that in bad weather the movement of both sexes was checked.

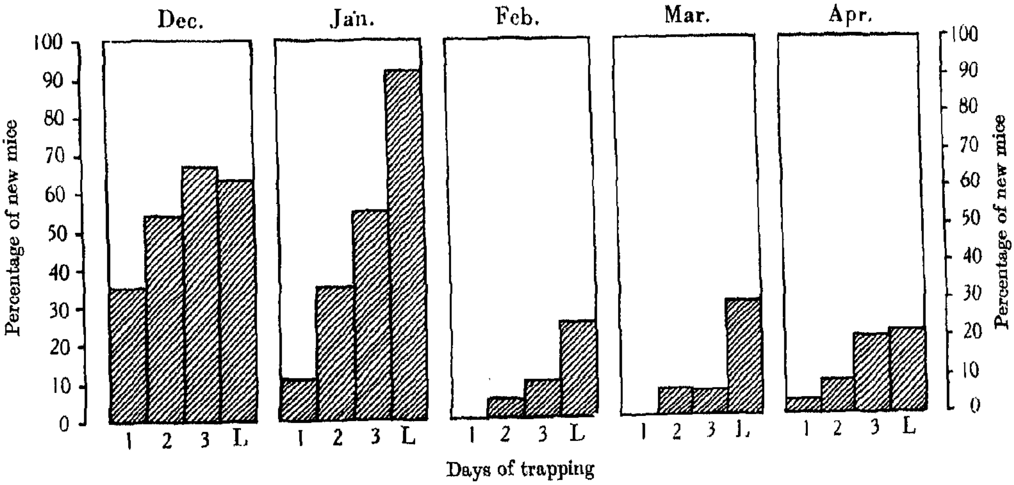
To summarize the data of Table 15, males and females and the two areas can be combined, and it will be found that those caught previously tend to be caught before the newcomers. In January only eight new mice were caught on the first day in spite of the fact that this month showed the largest number, seventy, of first day catches. In February and March no new mice were caught on the first day and in April only two out of the large catch

of fifty-eight. This tendency is demonstrated in Table 16 and Fig. 4 where the proportion of new mice in each day's catch is given as a percentage. Only percentages are given, as the actual numbers can easily be obtained from Table 15.

Still further information is gained by combining the records from each trapping site as in § 14. In that section the mice were grouped according to how often they were caught on any one site. Instead, they may be grouped according to the month in which they were

Table 16

Month	Percentage of new mice in each day's catch			
	First day	Second day	Third day	Later
Dec.	35	54	67	64
Jan.	11	35	55	92
Feb.	0	5	10	25
Mar.	0	7·5	7	30
Apr.	3·4	9	20	22



Note. L = fourth day and later.

Fig. 4. Percentage of new mice on each day of trapping (data from Table 16).

first caught. In Table 17 the mean day of catching of each such group is given for each month, the monthly arrays being kept separate because of the independent effect of season already pointed out on p. 350.

If the vertical monthly arrays are traced downwards in the table it will be seen that the mean day on which mice were caught in any month tends to become later as the month of first catching becomes later. As the groups are often small and many unknown factors must have affected the catching it is understandable that the results are not completely regular; the groups caught after February are very small indeed and any mean based on

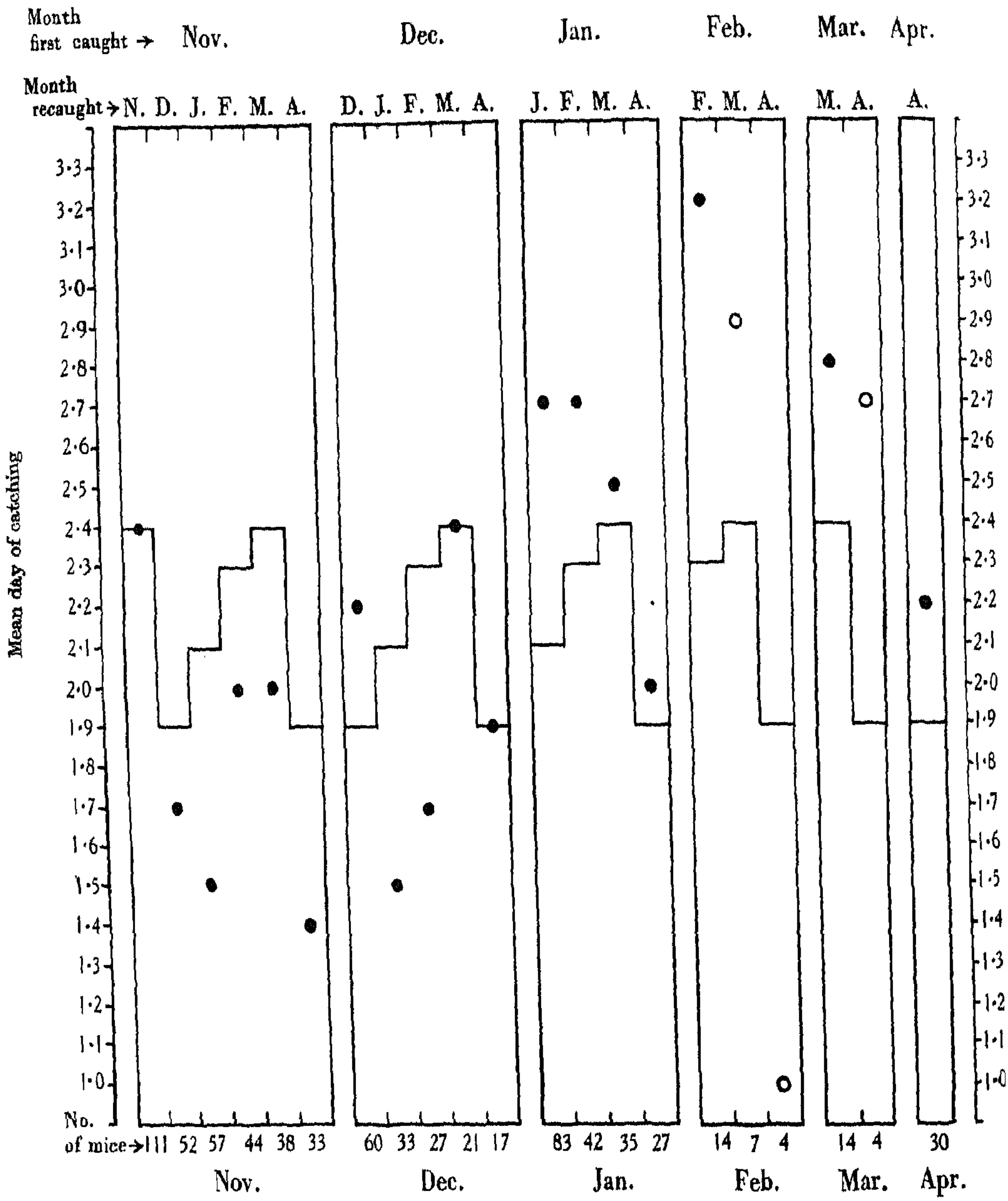


Fig. 5. Diagrammatic representation of the data of Table 17. The mean day of catching for all mice in each month, given at the bottom of Table 17, is shown by the horizontal step; these mean lines are the same in each rectangle. The dots represent the data from the body of that table and show that the mean day of catching tends to become later as the month in which the mouse was first caught becomes later. Where the mean depends on less than ten mice the result is shown by a circle instead of a dot.

less than eight mice is shown in italics to denote its unreliability. Nevertheless, the general tendency is clearly evident: the later the month in which a mouse was first caught the later the day of catching. Fig. 5 shows this tendency graphically by the method used in Fig. 3. The means from the arrays in Table 17 are shown as spots and the number of mice on which each is based is given below it; results depending on less than eight mice are represented by rings instead of spots. The line across each rectangle marks the monthly mean as before.

Table 17. *The mean day of catching for each month, for mice grouped according to the month in which they were first caught*

Month first caught	Month in which recaught					
	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.
Nov.	2·4	1·7	1·5	2·0	2·0	1·4
Dec.	—	2·2	1·5	1·7	2·4	1·9
Jan.	—	—	2·7	2·7	2·5	2·0
Feb.	—	—	—	3·2	*2·9*	*1·0*
Mar.	—	—	—	—	2·8	*2·7*
Apr.	—	—	—	—	—	2·2
Mean for each month	2·4	1·9	2·1	2·3	2·4	1·9

N.B. The numbers on which the means are based are shown at the bottom of Fig. 5. The figures in italics are means based on less than ten observations.

Since it has been shown that the day of catching is a good indication of how far a mouse lived from the traps, it seems that in each succeeding month mice living further and further afield were drawn into the traps. Also the young mice, as they grew in size, probably were able to wander further, just as it has been shown that males wander further than females. The fact that the new mice appeared on a late day shows that they were not immigrants settling in the place of those that disappeared. It seems that when there were many mice at the beginning of the winter the traps caught the mice from only a limited distance; as the season progressed and the mice became fewer they were caught from a wider area, the number caught each month remaining about the same.

17. SUMMARY

1. A description is given of a system of marking *Apodemus* by punching small holes in the ear pinnae.
2. The traps and nest boxes are described.
3. Our arrangement of traps in an area of Holwood Park, Keston, is described and our reasons are given for hoping to drain the area of mice. The number of days during which the traps were left out with this end in view is recorded. The effects of weather are discussed.

4. In an analysis of the efficiency of these methods it is shown that there was at least a 20 to 1 chance of a mouse being caught unless weather conditions were exceptionally unfavourable.

5. From the records of all mice caught in more than one month from December 1938 to April 1939, the proportion of mice surviving over each of the four trapping intervals is calculated. These four proportions are shown to approximate to a monthly survival rate of 0·876, or seven out of eight of the population.

6. The survival of two different series of mice is followed (1) from November 1938 to March 1940 and (2) from January 1939 to March 1940. The numbers surviving each month are compared with the numbers expected at the monthly survival rate of 0·876 calculated in § 7. After the commencement of the 1939 breeding season the survival rate of the winter population is shown to have been much reduced.

7. Further data are given on survival from one year to another. It appears that very few mice, in some years possibly none, survive from one winter season to the next.

8. Survival is analysed in relation to size and sex. In the winters 1938–40, a smaller proportion of the very small and very large mice appear to have survived than those of intermediate size. There was no appreciable difference in survival between the sexes.

9. The data on survival are discussed.

10. The survival of each month's catch of new mice is followed from November 1938 to March 1939 and the monthly survival ratios calculated. Reasons for the large proportion of mice caught once only are discussed.

11. Mice caught once only are shown to resemble mice caught in more than one locality in being caught on the average on a late day. It is therefore presumed that many of them lived at a distance from the traps.

12. It is shown that the less often a mouse was caught in any one locality the greater was its tendency to be caught on a late day; the day was also affected by the season but to a smaller extent.

13. The survival rate, unlike the day of catching, is shown not to change gradually with the number of times a mouse was caught, but to be uniquely low for first catches. This supports the evidence of the efficiency of trapping that this rate can be regarded as a true measure of survival and not merely of the failure of more distant mice to revisit the traps. The excessive number of single catches can be attributed to there being a limit to the distance from which mice could find their way home.

14. The replacement of mice by new arrivals is studied. When the mice are numerous at the beginning of winter the traps seem to catch mice from a limited distance; later, as the population becomes sparser and the young mice grow larger, they are caught from further afield.

We are indebted to Dr G. M. Morant for reading and criticizing the first draft of the paper and to Prof. E. S. Pearson for suggesting many improvements in the final stages. Miss Joyce Townend has again helped by preparing the diagrams.

REFERENCES

BARRETT-HAMILTON, G. E. H. & HINTON, M. A. C. (1910–21). *A History of British Mammals.* Gurney & Jackson, London.

BLAIR, W. F. (1941). Techniques for the study of mammal populations. *J. Mammal.* **22**, 148.

BLAIR, W. F. (1942). Size of home range and notes on the life history of the woodland deer mouse and eastern chipmunk in northern Michigan. *J. Mammal.* **23**, 27.

BOLE, B. P. (Jr.) (1939). The quadrat method of studying small mammal populations. *Sci. Publ. Cleveland Mus. Nat. Hist.* **5**, 15.

BURT, W. H. (1940). Territorial behaviour and populations of some small mammals in southern Michigan. *Misc. Publ. Mus. Zool. Univ. Mich.* no. 45.

CHITTY, D. (1937). A ringing technique for small mammals. *J. Anim. Ecol.* **6**, 36.

DIVER, C. (1933). The physiography of South Haven Peninsula, Studland Heath, Dorset. *Geogr. J.* **81**, 404.

DIVER, C. & GOOD, R. D'O. (1934). The South Haven Peninsula survey (Studland Heath, Dorset): general scheme of the survey. *J. Anim. Ecol.* **3**, 129.

ELTON, C., FORD, E. B., BAKER, J. R. & GARDNER, A. D. (1931). The health and parasites of a wild mouse population. *Proc. Zool. Soc. Lond.* p. 657.

ELTON, C. (1942). *Voles, Mice and Lemmings.* Clarendon Press, Oxford.

EVANS, F. C. (1942). Studies of a small mammal population in Bagley Wood, Berkshire. *J. Anim. Ecol.* **11**, 182.

HACKER, H. P. & PEARSON, H. S. (1944). The growth, survival, wandering and variation of the long-tailed field mouse, *Apodemus sylvaticus.* I. Growth. *Biometrika*, **33**, 136.

SUMNER, F. B. & COLLINS, H. H. (1918). Autotomy of the tail in rodents. *Biol. Bull. Woods Hole*, **34**, 1.

TABLE OF PERCENTAGE POINTS OF THE *t*-DISTRIBUTION

By ELIZABETH M. BALDWIN, *Post Office Research Station*

In the application of tests of significance a need is sometimes felt for a table of the 5 and 1% points of the t-distribution when the number of degrees of freedom n is greater than 30. It was found that the use of the normal probability curve, taking $t\sqrt{\frac{n-2}{n}}$ as a normal deviate (as recommended in some text-books, e.g. Rider, 1939, p. 89), gave results much smaller than the true values. This note contains a table of percentage points of the t-distribution which has been calculated to cover this need.

Percentage points of t

n (no. of degrees of freedom)	t 95%	t 99%	n (no. of degrees of freedom)	t 95%	t 99%	n (no. of degrees of freedom)	t 95%	t 99%
1	12·706	63·657	23	2·069	2·807	58	2·001	2·663
2	4·303	9·925	24	2·064	2·797	60	2·000	2·660
3	3·182	5·841	25	2·060	2·787	62	1·999	2·658
4	2·776	4·604	26	2·056	2·779	64	1·998	2·655
5	2·571	4·032	27	2·052	2·771	66	1·996	2·652
6	2·447	3·707	28	2·048	2·763	68	1·995	2·650
7	2·365	3·499	29	2·045	2·756	70	1·994	2·648
8	2·306	3·355	30	2·042	2·750	72	1·993	2·646
9	2·262	3·250				74	1·992	2·644
10	2·228	3·169	32	2·037	2·738	76	1·992	2·642
11	2·201	3·106	34	2·032	2·728	78	1·990	2·640
12	2·179	3·055	36	2·028	2·720	80	1·989	2·639
13	2·160	3·012	38	2·024	2·712	82	1·988	2·637
14	2·145	2·977	40	2·021	2·704	84	1·987	2·635
15	2·131	2·947	42	2·013	2·698	86	1·987	2·634
16	2·120	2·921	44	2·015	2·692	88	1·986	2·632
17	2·110	2·898	46	2·013	2·687	90	1·986	2·631
18	2·101	2·878	48	2·010	2·682	92	1·986	2·630
19	2·093	2·861	50	2·008	2·678	94	1·986	2·629
20	2·086	2·845	52	2·006	2·674	96	1·984	2·627
21	2·080	2·831	54	2·005	2·670	98	1·983	2·626
22	2·074	2·819	56	2·003	2·667	100	1·982	2·625

In computing the values of t given in the table, use was made of *Tables of Percentage Points of the Incomplete Beta Function* (Thompson, 1941) to give a first approximation. The final values were then obtained by interpolation (using the trivariate Everett formula) from Pearson's *Tables of the Incomplete Beta Function* (Pearson, 1934) and should be correct to the three decimal places given.

It should be noted that this table is an extension, for the 5 and 1% probability levels, of the table computed by Maxine Merrington (1942).

REFERENCES

Merrington, Maxine (1942). *Biometrika*, **32**, 300.

Pearson, Karl (1934). *Tables of the Incomplete Beta Function*, **1**. Cambridge University Press.

Rider, P. R. (1939). *Modern Statistical Methods*. John Wiley and Sons, Inc., London: Chapman and Hall.

Thompson, Catherine M. (1941). *Biometrika*, **32**, 168.